Yu M. KORSHUNOV

FUNDAMENTOS MATEMATICOS de la CIBERNETICA


ИЗДАТЕЛЬСТВО
МИР

Ю. М. КОРШУНОВ

МАТЕМАТИЧЕСКИЕ ОСНОВЫ КИБЕРНЕТИКИ

ИЗДАТЕЛЬСТВО „ЭНЕРГИЯ“
МОСКВА

Yu. M. KORSHUNOV

FUNDAMENTOS MATEMATICOS de la CIBERNETICA

EDITORIAL MIR
MOSCÚ

TRADUCIDO DEL RUSO
POR EL INGENIERO
ROBERTO LANIER
ALVAREZ

Impreso en la URSS

На испанском языке

INDICE

PARTE PRIMERA

FUNDAMENTOS DE MATEMATICAS DISCRETAS

CAPÍTULO QUINTO. CORRELACIONES DE LA TEORÍA DE LOS CONJUNTOS EN LA TEORÍA DE PROBABILIDADES Y ELEMENTOS DE ESTADÍSTICA MATEMATICA 134

PARTE SEGUNDA

OPTIMIZACION DE LOS PROCESOS DE MANDO

PRÓLOGO

Las nuevas corrientes científicas surgidas en los últimos años que se abarcan con la palabra "cibernética", encuentran su aplicación más vasta en la rama de la automática y telemecánica. En las disciplinas explicadas en esta especialidad se reflejan las cuestiones de la teoría de la información, teoría de la detección estadística de señales, teoría de los sistemas óptimos, teoría de los dispositivos automáticos terminales, así como una serie de principios basados en las teorías de los juegos y de las decisiones estadísticas.

Sin embargo, la exposición de los asuntos de cibernética en cursos especiales tropieza con la seria dificultad de que el curso tradicional de matemáticas superiores de los centros de enseñanza superior técnica, que hace hincapié en lo continuo y determinado, no brinda el fundamento requerido para el estudio de estas nuevas disciplinas cuya base descansa en lo discreto y aleatorio. Estos asuntos pueden exponerse en las disciplinas "Fundamentos Matemáticos de la Cibernética" o "Fundamentos Teóricos de la Cibernética". No obstante, en la actualidad no hay libros de texto algunos sobre estas asignaturas.

La presente obra constituye el primer intento de subsanar esta deficiencia existente. Ella ha sido preparada fundamentándose en la enseñanza de dicha disciplina a los estudiantes de la especialidad de Automática y Telemecánica y de otras especialidades afines en el Instituto Radiotécnico de Riazán.

El libro consta de la introducción y de dos partes que comprenden diez capítulos.

En la introducción, además del concepto de sistema cibernético, se brindan algunos datos sobre los sistemas de comunicación y control, lo que permite ilustrar ciertos métodos matemáticos con ejemplos afines a la especialidad.

En la primera parte se exponen varios capítulos de las matemáticas muy necesarios al ingeniero pero no incluidos en el programa de matemáticas superiores de los centros de enseñanza superior técnica. El material de la primera parte tiene gran valor independiente, ya que permite resolver una serie de tareas de gran importancia práctica en la esfera de la automática y telemecánica.

No obstante, la elección del contenido se ha realizado de tal modo que la primera parte sirve de introducción matemática a la segunda parte dedicada a la optimización de los procesos de control y que tiene por objetivo introducir al estudiante en el círculo de conceptos de la teoría de control moderna.

Por supuesto que en un libro de volumen limitado como éste sólo fue posible reflejar una pequeña parte de los conceptos y métodos utilizados en cibernética. Esto se refiere tanto a la primera parte de la obra como, en particular, a la segunda. En especial, la optimización de los procesos de control que viene a ser parte esencial de la teoría general del control, se basa en gran medida, en el análisis matemático y el cálculo de variaciones clásicos. Los métodos de optimización del control elaborados sobre dicha base, estudiados por lo común en el curso "Teoría del control automático" y en algunos cursos especiales, no han sido examinados detalladamente en las páginas de la obra presente. La atención principal se ha prestado a aquellas ideas y métodos que surgieron gracias al desarrollo de la cibernética y la aparición de las calculadoras numéricas. Por otra parte, también en estas cuestiones el autor, en varios casos, se ha visto obligado a limitarse sólo al bosquejo esquemático de las ideas fundamentales.

En necesario señalar también que aunque casi todos los métodos estudiados se ilustran con ejemplos y problemas, los métodos numéricos de cálculo sólo pueden dominarse con éxito a base de la resolución independiente de problemas, lo que puede garantizarse realizando clases prácticas y distribuyendo tareas de cálculo sobre este curso.

El autor considera un deber expresar su gratitud a los críticos L. T. Kuzin y D. A. Pospélov, así como al redactor N. I. Glazunov cuyas valiosas observaciones contribuyeron a mejorar el contenido de la obra.

El autor

INTRODUCCIÓN

1-1. OBJETO DE LA CIBERNETICA

La cibernética es una ciencia joven que surgió en los primeros años después de la Segunda Guerra Mundial y se desarrolló tan impetuosamente que en la actualidad ha conquistado posiciones firmes en muchas ramas de la ciencia y de la técnica. La cibernética debe sus éxitos al descubrimiento de una serie de analogías entre el funcionamiento de los dispositivos técnicos, la actividad vital de los organismos y el desarrollo de las colectividades de seres vivos. La cibernética reforzó estas analogías derivadas de razonamientos generales de carácter metodológico, creando métodos matemáticos que permitían describir desde un punto de vista cuantitativo, los procesos que ocurren en sistemas de la naturaleza física más diversa. Los principios de la cibernética encuentran vasta aplicación en automática y telemecánica, teoría de la comunicación, economía, sociología, biología y medicina.

La propia palabra "cibernética" es de origen griego. Los antiguos griegos designaban con esta palabra el arte de gobernar las naves. En el siglo XVIII el término "cibernética" se utiliza por el eminente físico y matemático A. M. Ampère para definir la ciencia de gobernar el Estado. En su interpretación contemporánea se entiende por cibernética la ciencia del control en el sentido más amplio de esta palabra. La acepción presente del término "cibernética" está vinclulada al nombre del gran matemático norteamericano N. Wiener, cuya obra "Cibernética o control y comunicación en el Animal y la Máquina", que vio la luz en 1948, dio inicio a la formación de esta nueva disciplina científica.

La aparición de la cibernética como ciencia del control, está ligada al progreso técnico general que caracteriza el desarrollo de las fuerzas productivas en la época actual.

Antes de surgir la cibernética, las principales direcciones del desarrollo de la técnica se caracterizaban, en primer lugar, por la creación de dispositivos que servían para obtener y transformar la energía (por ejemplo, máquinas de vapor, turbinas, generadores de energía eléctrica, motores eléctricos y de otros tipos, etc.) y, en segundo, por la creación de dispositivos destinados a influir sobre

la naturaleza circundante. En tales dispositivos la atención principal está puesta en las relaciones energéticas siendo el rendimiento el índice más importante de su funcionamiento. La sencillez relativa de los dispositivos técnicos no situaba en un lugar particular el problema del control de los mismos. El hombre trabajaba y mandaba simultáneamente el objeto de su trabajo. La información requerida para el mando la obtenía directamente de sus órganos sensitivos, observando los resultados de su trabajo.

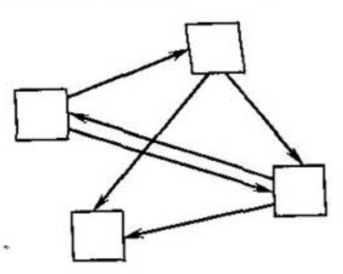
Sin embargo, el progreso de la técnica condujo a mediados del siglo XX a la creación de sistemas técnicos tan complejos que los problemas del control de los mismos comenzaron a superar las posibilidades fisiológicas del ser humano. A finales de la Segunda Guerra Mundial tal problema fue la creación de un sistema de mando automático del fuego antiaéreo, que, siendo las velocidades de los aviones comparables a la velocidad del proyectil antiaéreo, podría seguir el curso de los aviones, realizar el cálculo de sus trayectorias y apuntar los cañones sin la intervención del hombre. En tales sistemas se promueven a primer plano los problemas de la obtención de información sobre las condiciones ambientales y de la elaboración de dicha información con el objeto de extraer de ella datos útiles para el control y del empleo de dicha información para realizar actividades orientadas hacia una finalidad determinada, o sea, los problemas de la creación de dispositivos que sirvan para la *comunicación* y el *control*. La necesidad de resolver estos problemas provocó un rápido progreso en la esfera de la teoría de la comunicación, técnica de cálculo y automática, lo que dio inicio al desarrollo de aquellas ideas que posteriormente constituyeron el fundamento de la cibernética.

Los dispositivos de comunicación y control difieren substancialmente de los dispositivos técnicos apuntados más arriba, en el sentido de que en éstos, las relaciones energéticas no juegan un papel primordial y se presta atención principal a su capacidad de transmitir y elaborar sin distorsiones grandes cantidades de información. Así, en una línea de radiocomunicación, sólo llega al receptor una parte ínfima de la energía irradiada por la antena del radiotransmisor y el rendimiento resulta sumamente bajo. No obstante, la línea de radiocomunicación se considera buena si los mensajes se transmiten por la misma con pequeñas distorsiones y no están sometidos a la influencia de las interferencias. De este modo, los procesos fundamentales en los dispositivos de comunicación y de control, son los de transmisión y elaboración de la información y no los procesos relacionados con la transformación y utilización de la energía.

La carencia de importancia de las relaciones energéticas en los problemas de comunicación y de control permite abstraerse de las particularidades físicas de los portadores de la información y de la naturaleza física de los sistemas en los cuales se utiliza dicha información. Por eso la cibernética es la teoría general de la comu-

nicación y el control, aplicable a cualquier sistema, independientemente de su naturaleza física.

El concepto de *sistema*, además del concepto de *control* cuyo sentido exacto será aclarado más adelante, es el concepto fundamental de la cibernética. Cualquier sistema real existente consta de objetos concretos, en calidad de los cuales pueden intervenir dispositivos técnicos, individuos que mandan dichos dispositivos, recursos materiales, etc. Estos objetos están vinculados entre sí y con el mundo circundante mediante determinados enlaces que constituyen fuerzas y flujos de energía, materia e información. Pero la cibernética prescinde del contenido físico de las propiedades de los objetos y enlaces considerando el sistema real como un conjunto abstracto de elementos dotados de propiedades comunes que se encuentran en ciertas relaciones mutuas determinadas por el carácter de los enlaces existentes. Tal concepción permite renunciar a la división acostumbrada de los sistemas en mecánicos, eléctricos, químicos, biológicos, etc., e introducir el concepto de sistema cibernético abstracto como conjunto de elementos interrelacionados que influyen unos sobre otros. En la figura I-1 se da un ejemplo de sistema cibernético que consta de cuatro elementos y seis enlaces mutuos.


Fig. I-1. Ejemplo de un sistema cibernético

El examen del sistema como conjunto de elementos posibilita la utilización de la teoría de los conjuntos como instrumento para su descripción matemática. Por cierto, en algunos casos de importancia es conveniente emplear la lógica matemática como instrumento para describir los enlaces entre los elementos. Por eso, la teoría de los conjuntos y la lógica matemática, con cuya exposición se inicia la presente obra, constituyen la base de la descripción matemática de un sistema.

Los sistemas que se encuentran en la práctica se dividen, atendiendo a la estructura y el carácter de sus enlaces, en determinados y probabilísticos. Se denomina determinado el sistema cuyas leyes de movimiento se conocen con exactitud y su comportamiento futuro puede preverse. Para un sistema probabilístico es imposible hacer un pronóstico preciso de su comportamiento futuro. El mecanismo de la relojería puede servir de ejemplo de un sistema determinado. Sin embargo, los sistemas de control estadístico de la producción, los sistemas de arribo de buques a los puertos marítimos o la reserva de mercancías en un almacen que tiene un gran número de suministradores y consumidores son sistemas probabilísticos.

Los problemas que resuelve la cibernética conducen en la mayoría de los casos a la necesidad de examinar sistemas probabilísticos bastante complejos que constan de un gran número de elementos y tienen enlaces internos variados y ramificados. Precisamente a tales sistemas pertenecen la mayoría de los sistemas industriales, los sistemas económicos, sociales y biológicos. Para la descripción matemática de tales sistemas, además de la teoría de los conjuntos y la lógica matemática, se utilizan ampliamente como instrumento la teoría de las probabilidades y los métodos de estadística matemática.

Hasta ahora sólo nos hemos referido a los métodos matemáticos empleados para describir sistemas cibernéticos. Pero el objetivo de la cibernética es el control de los sistemas. Para juzgar sobre las vías de solución de tal tarea es necesario comprender con claridad el sentido del término "control".

En el amplio sentido de la palabra, se entiende por *control* la actividad organizativa encaminada a la consecución de objetivos determinados que ejecuta las funciones de dirección del trabajo ajeno. El proceso de control consiste en la toma de decisiones sobre las acciones más convenientes en una u otra situación creada. La persona que lleva a cabo el control toma sus decisiones evaluando las condiciones ambientales con ayuda de la información recibida de sus órganos sensitivos, los instrumentos de medida u otras personas. En muchos casos esta información resulta insuficiente para la estimación unívoca de las condiciones ambientales. Entonces el hombre utiliza su experiencia, conocimientos, memoria e intuición. Es un excelente rasgo característico del ser humano, su capacidad para tomar decisiones en situaciones de una considerable indeterminación en lo que respecta a las circunstancias ambientales.

No obstante, en las condiciones de las grandes empresas industriales modernas, incluso los conocimientos y la intuición de un dirigente experimentado resultan insuficientes para realizar un control eficaz. Como resultado surgen deficiencias en el funcionamiento de las grandes empresas tales como el trabajo "a la carrera", dificultades con el suministro regular de materias primas y materiales sin el aumento excesivo de las reservas, problemas de transporte, etc.

La cibernética se plantea la tarea de facilitar al hombre el proceso de la toma de decisiones de importancia, encomendando a los dispositivos automáticos la recogida y elaboración de grandes cantidades de información respecto al estado del proceso de producción, el análisis de las situaciones creadas y la elaboración de recomendaciones acerca de las acciones dirigidas a cierta finalidad. Los dispositivos automáticos que ejecutan el conjunto de tales operaciones se denominan sistemas automáticos de control. El funcionamiento de tales sistemas se basa en el empleo de las calculadoras electrónicas numéricas (C. E. N.).

El papel de las C. E. N. en la cibernética resulta tan importante que es menester detenerse para estudiar más detalladamente esta cuestión.

Inicialmente las C. E. N. se utilizaban para realizar cálculos tradicionales que antes ocupaban muchas horas y ahora sólo requerían segundos. Pero pronto se hizo evidente que enorme aumento de la velocidad de los cálculos entrañaba efectos cualitativamente nuevos. Si antes el proyectista o el economista sólo podía analizar algunas, de las variantes de solución, que por ciertos motivos le parecieron dignas de atención, ahora se presentaba la oportunidad de comparar todas las variantes posibles y elegir la mejor de todas. Así surgieron las ideas de la optimización que posteriormente condujeron al desarrollo de una serie de nuevos capítulos de las matemáticas.

Luego resultó que la C. E. N. instalada en una empresa industrial es capaz de elaborar con facilidad grandes cantidades de información sobre la marcha del proceso de producción y puede convertirse en ayudante insustituible del hombre en la administración de la producción.

No obstante, para que las C. E. N. puedan ser utilizadas con fines de control, deben ser elaborados métodos matemáticos que permitan analizar los tipos de información existentes, desechar la información superflua y separar su parte más substancial, empleando esta información para estimar las situaciones creadas y preparar recomendaciones que garanticen el cumplimiento más efectivo de los objetivos del control. La necesidad de resolver problemas análogos provocó la aparición de tales capítulos de las matemáticas como la teoría de la información, teoría de los juegos, teoría de las decisiones estadísticas, teoría de colas, la programación lineal y dinámica y otros. En la presente obra se estudiará una parte de estos nuevos métodos matemáticos. Los demás métodos se estudian en diferentes cursos especiales.

1-2. TRANSMISION Y CODIFICACION DE LA INFORMACION

Los enlaces entre los elementos de cierto sistema pueden servir para distintos objetivos. Por los mismos pueden transmitirse energía, materia, esfuerzos, etc. Pero en los sistemas cibernéticos nos interesa, en primer lugar, el contenido informativo de los enlaces, es decir, la posibilidad de utilizar dichos enlaces para transmitir datos sobre los diversos estados de los elementos del sistema.

En técnica, cualquier dato sobre algún suceso ocurrido dentro o fuera del sistema se denomina *mensaje*. Los enlaces informativos que sirven para transmitir los mensajes se denominan *canales de comunicación*. Los portadores físicos de la información en los canales de comunicación se llaman *señales*. En la figura 1-2 se

muestra un esquema estructural que representa un canal de comunicación.

La palabra "señal" procede de la voz latina *signum*, o sea, que denota símbolos convencionales eléctricos, acústicos, o de otra naturaleza que sirven para transmitir mensajes. El tipo de las señales es, por lo general, el resultado de un acuerdo entre el hombre que transmite la información y el que la recibe y no guarda relación directa con el contenido de la información transmitida. Por eso las señales pueden convertirse fácilmente de un tipo a otro sin alterar su contenido informativo y también pueden almacenarse (por ejemplo, en forma de registro literal) para utilizar en el futuro dicho contenido.

Por ejemplo, examinemos el tipo de señales empleadas al transmitir un telegrama. El mensaje inicial impreso en el formulario telegráfico, se convierte a la forma del código Morse que

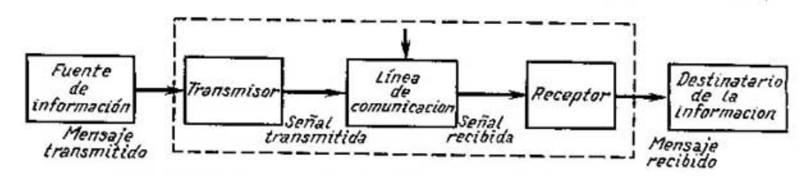


Fig. I-2. Esquema estructural de un canal de comunicación

consta de puntos y rayas transmitidos mediante impulsos de corriente, largos y cortos, por la línea de comunicación. En la parte de recepción, por los impulsos captados, se regenera el texto inicial del mensaje. Como vemos, en este caso el mensaje existe en calidad de señales de diferente forma: de texto literal, puntos y rayas del código Morse, impulsos de corriente en la línea de comunicación, etc. La conversión de una señal de una forma a otra se denomina *codificación*.

El método para realizar una comunicación puede representarse del modo siguiente. Ante todo debe existir un grupo de símbolos: letras, palabras, puntos, rayas, etc., que tengan un sentido conocido tanto por el remitente como por el destinatario del mensaje. El grupo de tales símbolos se denomina *alfabeto*. Los propios símbolos se determinan por convenio de las partes.

La teoría de la comunicación se basa en el postulado de que los símbolos componentes de alfabeto no pueden ser infinitamente diversos. Por eso, para la transmisión de mensajes cualesquiera se utiliza un número limitado de símbolos diferentes. Así, todos los mensajes literales posibles en lengua rusa se componen con ayuda de un alfabeto de 33 letras.

Durante el proceso de transmisión el remitente elige del alfabeto existente un símbolo tras otro, los convierte en las señales

correspondientes y las transmite por el canal de comunicación. En el canal de comunicación, las señales se someten a la influencia de las interferencias, lo que provoca su distorsión. Así, las señales en la parte de recepción van a diferir de las señales enviadas al canal de comunicación.

El proceso de recepción consiste en que el destinatario, al recibir cualquier señal, debe identificarla con uno de los símbolos existentes del alfabeto, o sea, debe excluir todos los demás símbolos, salvo uno. Esta tarea puede presentar considerables dificultades si en el canal de comunicación las señales se someten a grandes distorsiones. Los métodos para superar estas dificultades constituyen la esencia de la teoría de la comunicación.

En los sistemas técnicos de comunicación se emplean alfabetos de varios tipos. Sin embargo, por diversos motivos, tiene gran empleo el *alfabeto binario* que sólo utiliza dos tipos de símbolos designados convencionalmente por 0 y 1. En el alfabeto binario cualquier mensaje va a constituir una sucesión de ceros y unidades, por ejemplo 100110100.

Es fácil calcular que el número total de mensajes que constan de m letras del alfabeto binario, será igual a 2^m. En particular, cualquier letra del alfabeto puede representarse por seis signos del alfabeto binario, por ejemplo: a, 000001; b, 000010; c, 000011, etc. Como que seis signos binarios proporcionan $2^6 = 64$ diferentes combinaciones de símbolos, con su auxilio pueden representarse no sólo todas las letras del alfabeto, sino también los signos de puntuación. Por consiguiente, con ayuda del alfabeto binario puede representarse y transmitirse por un canal de comunicación cualquier mensaje literal.

El alfabeto binario puede utilizarse asimismo para transmitir datos numéricos, pero en tal caso resulta necesario emplear ciertos sistemas de numeración especiales.

En el sistema de numeración decimal ampliamente utilizado, los diversos números se escriben con ayuda de diez cifras, 0, 1, ..., 9, dispuestas en determinado orden y que tienen valores dependientes de la ubicación de cada cifra. Así, la inscripción 395 representa el número definido por la expresión

$$3 \cdot 10^2 + 9 \cdot 10^1 + 5 \cdot 10^0.$$

Aquí, el número 10 se denomina *base* del sistema de numeración.

Análogamente, cualquier número N puede escribirse en un sistema de numeración de cualquier otra base R (número entero), con ayuda de distintas cifras cuyo número es igual a la base del sistema de numeración. En tal caso la inscripción $\ldots d_3 d_2 d_1 d_0$, donde d_i son las cifras del número $N (0 \leqslant d_i < R)$, define la magnitud

$$N = \ldots d_3R^3 + d_2R^2 + d_1R^1 + d_0R^0. \qquad \text{(I-1)}$$

Así, en el sistema de numeración óctuple que se emplea en ciertos tipos de calculadoras numéricas, el número 395 tendrá el aspecto

$$6 \cdot 8^2 + 1 \cdot 8^1 + 3 \cdot 8^0,$$

es decir, se escribirá en forma del número 613.

En caso de utilizarse el alfabeto binario, la inscripción del número deberá efectuarse sólo mediante las cifras 0 y 1. Sirve para dicha notación el sistema de numeración binario cuya base es el número 2. Cualquier número del 0 al 15 puede representarse en el sistema de numeración binario por medio de un número de cuatro dígitos:

$$3 = 0 \cdot 2^3 + 0 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0, \text{ o sea, } 0011;$$

$$5 = 0 \cdot 2^3 + 1 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0, \text{ o sea, } 0101;$$

$$9 = 1 \cdot 2^3 + 0 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0, \text{ o sea, } 1001.$$

Naturalmente, no es imprescindible escribir los ceros de órdenes superiores de estos números, es decir, los números 3 y 5 pueden escribirse en forma de 11 y 101. El número 395, que puede ser representado por medio de la base 2:

$$395 = 2^8 + 2^7 + 2^3 + 2^1 + 2^0,$$

se escribirá en el sistema de numeración binario con un número de nueve dígitos 110001011.

La anotación de números grandes en el sistema de numeración binario presenta el inconveniente porque requiere una gran cantidad de dígitos, lo que dificulta la lectura de los números y la estimación rápida de su magnitud. Es por ello que con frecuencia se emplean sistemas de numeración mixtos, por ejemplo, el decimal-binario, en el cual el propio número se escribe en el sistema de numeración decimal y sus diferentes dígitos, en el sistema de numeración binario, utilizando cuatro dígitos binarios para cada cifra decimal. De este modo, en el sistema de numeración decimal-binario, el número 395 tiene el aspecto:

$$\underbrace{0011}_{3} \quad \underbrace{1001}_{9} \quad \underbrace{0101}_{5}.$$

1-3. CONCEPTO DE SISTEMA CONTROLADO

Como se desprende de la definición del concepto de control, dada más arriba, cualquier sistema controlado puede representarse en forma de un conjunto de dos partes: la parte controlada denominada también *objeto* controlado, y la parte que controla llamada *dispositivo de control* u *operador*.

Cada objeto controlado se caracteriza por determinadas propiedades:

1) la existencia de una determinada designación orientada hacia una finalidad que se manifiesta en la capacidad de dar cualquier resultado beneficioso;

2) el estado del objeto que se manifiesta en tipos concretos de movimiento y puede cambiar al variar las condiciones ambientales en las cuales se halla el objeto;

3) la capacidad de ser controlable, es decir, la capacidad del objeto de reaccionar a las influencias externas ejercidas sobre sus órganos especiales, los órganos de control.

La tarea del operador consiste en garantizar el cumplimiento por el objeto de su finalidad actuando sobre los órganos especiales, los órganos de control. Para ello debe ocurrir un intercambio de información entre el operador y el objeto controlado. Constituyen tal información, por un lado, las señales de mando con cuya ayuda el operador influye sobre el objeto controlado, o sea, la *información de mando* y, por otro lado, los datos sobre el estado del objeto

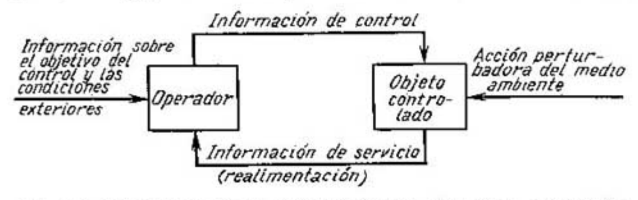


Fig. I-3. Circulación de la información en el sistema controlado

controlado a base de los cuales el operador determina el tipo de señales de mando, es decir, la *información de servicio.*

En la figura I-3 se muestra el esquema estructural que presenta los flujos de información principales en el sistema controlado. En dicho esquema, además de los flujos de circulación de las informaciones de mando y de servicio dentro del sistema, se presenta el enlace del objeto controlado y del operador con el medio ambiente. La variación de las condiciones ambientales ejerce una influencia inmediata sobre el objeto controlado provocando el cambio del carácter de su movimiento y dificultando el cumplimiento de su finalidad. Para controlar con acierto el objeto, el operador debe recibir información sobre las condiciones ambientales, tomarla en consideración al elaborar las señales de mando y disponer también de información sobre la finalidad del control.

La información de mando se obtiene elaborando todo tipo de datos que llegan al operador. Una parte de la información puede almacenarse para ser utilizada posteriormente. Puede ejecutar las funciones de operador una persona o un dispositivo mecánico o electrónico. En los últimos tiempos es frecuente que las calculadoras electrónicas universales o especializadas jueguen el papel de operador.

En el objeto controlado ocurre el proceso de elaboración de la información de mando que se manifiesta en el cambio del carácter del movimiento de dicho objeto. Estos cambios se transmiten por el canal de realimentación al operador en forma de información de servicio.

La *realimentación* juega un papel importantísimo en la realización de un mando eficaz, ya que brinda al operador la posibilidad de juzgar constantemente en el proceso de control, en qué medida se ha logrado la finalidad del control, y de acuerdo con esto, elaborar de modo más racional las señales de mando. Por eso, el principio de realimentación es el fundamento de la inmensa mayoría de los procesos de control, y en particular, la base de casi toda la actividad del ser humano. Por ejemplo, si una persona extiende la mano para tomar de la mesa un lápiz, ella realiza de modo inconsciente e ininterrumpido la comparación de la situación mutua de la mano y el lápiz, gracias a lo cual el movimiento resulta muy preciso.

El objeto del curso "Teoría del control automático" es el estudio detallado del principio de realimentación. La tarea del curso "Fundamentos Matemáticos de la Cibernética" es el análisis de los métodos de descripción matemática de los sistemas de control automáticos y la explicación de los procedimientos y métodos de toma de decisión en los cuales se basa el diseño de dispositivos de control.

Aclararemos el esquema general del proceso de control mediante ejemplos de sistemas de diversa naturaleza física.

Ejemplo I-1. El regulador centrífugo de la velocidad de una máquina de vapor (fig. I-4). En el sistema dado el regulador centrífugo de velocidad juega el papel de operador. El órgano de control es la mariposa que regula el paso del vapor a la máquina de vapor. Antes de comenzar el funcionamiento del

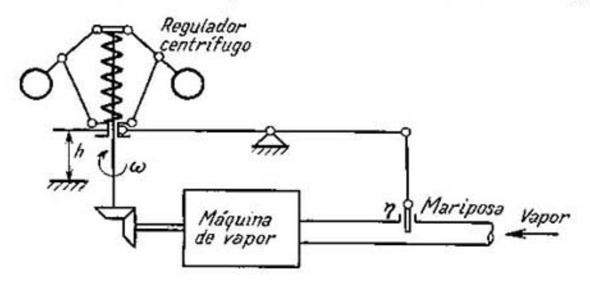


Fig. I-4. Esquema de regulación de la velocidad de la máquina de vapor

regulador debe introducirse en él información sorbe la finalidad del control lo que se logra ajustando el regulador para mantener una velocidad de rotación determinada ω_0 mediante la ubicación correspondiente de los pesos y el cambio de la tensión de los muelles. La información de servicio sobre la velocidad de rotación real de la máquina de vapor ω, se transforma por el re-

gulador centrífugo en cambio de posición del manguito h y dicha información de servicio se convierte por medio de una palanca en información acerca de la posición de la mariposa que acciona sobre el acceso del vapor a la máquina y cambia de este modo su velocidad de rotación. La influencia del medio ambiente se manifiesta aquí en forma de variación de la carga de la máquina de vapor, lo que provoca el cambio de su velocidad de rotación.

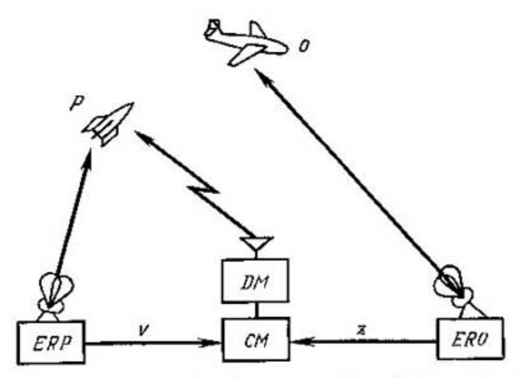


Fig. I-5. Sistema de mando de un proyectil antiaéreo teleguiado

Ejemplo I-2. El jefe de movimiento del transporte ferroviario. En el ejemplo dado, cumple la función de operador la persona jefe de movimiento que hallándose en el puesto de gobierno recibe por el selector o teléfono la información de servicio sobre el estado de los trenes en aquel sector de la línea del cual es responsable. Esta información consiste en datos acerca de la disposición para la partida de un tren de turno, el retraso de un tren con respecto al gráfico de marcha de los trenes, la necesidad de partida de un tren fuera de turno (no previsto en el gráfico), el grado de carga de la línea, etc. La obligación del jefe de movimiento consiste en transformar la información recibida y elaborar decisiones encaminadas al cumplimiento óptimo del gráfico de marcha de los trenes. Esta información de control consta de disposiciones para los maquinistas o jefes de estaciones sobre el orden ulterior del movimiento de los trenes (sobre la aceleración o deceleración de la marcha de los trenes, la maniobra del paso de los mismos a los apartaderos, etc.).

Ejemplo I-3. El mando del prouectil antiaéreo. En calidad de ejemplo del sistema automático en el cual juegan el papel de operador dispositivos calculadores electrónicos, en la figura I-5 se muestra el esquema funcional de mando de un proyectil antiaéreo teledirigido a un objetivo móvil. En este sistema, la información de servicio está contenida en las señales suministradas por la estación de radiodetección E. R. O., que determina las coordenadas del objetivo móvil O y por la estación de radiodetección E. R. P., que determina las coordenadas del proyectil P que se mueve hacia el objetivo. Basándose en estos tipos de información, la calculadora de mando elabora las señales que llegan por la línea de radio al dispositivo de mando del proyectil antiaéreo y que corrigen su movimiento según la posición mutua del proyectil y el objetivo.

PROBLEMAS PARA LA INTRODUCCION

I-1. Un sistema consta de siete elementos, cada uno de los cuales se comunica con todos los demás mediante enlaces del tipo más sencillo que tienen sólo dos estados: "hay enlace" y "no hay enlace". ¿Cuál es el número posible

de estados de tal sistema? Si es preciso indicar mediante un experimento todos los estados que poseen la propiedad P y para investigar cada estado se requiere 1 s, ¿cuánto tiempo se necesitará para investigar todos los estados posibles?

I-2. ¿En cuánto disminuye el número de estados posibles del sistema en el problema 1, si sólo hay enlaces direccionales sucesivos (fig. I-6, *a*) o el sistema tiene estructura jerárquica (fig. I-6, *b*)? Cite ejemplos de sistemas que tengan estructuras de ambos tipos.

I-3. El polinomio (I-1) que sirve para la representación de un número en el sistema de numeración de base R, puede escribirse en la forma

$$N = \{[(\ldots d_3) R + d_2] R + d_1\} R + d_0, \qquad (1\text{-}2)$$

lo que equivale a la secuencia de fórmulas

$$N = N_1R + d_0, \quad N_1 = N_2R + d_1, \quad N_2 = N_3R + d_2 \ \ldots$$

Fundamentar con ayuda de estas fórmulas la siguiente regla de conversión del número 395 de la forma decimal a la óctuple

$395 : 8 = 49$ resto 3 o más brevemente
$49 : 8 = 6$ resto 1
$6 : 8 = 0$ resto 6

395	
49	3
6	1
0	6,

lo que da el número óctuple 613.

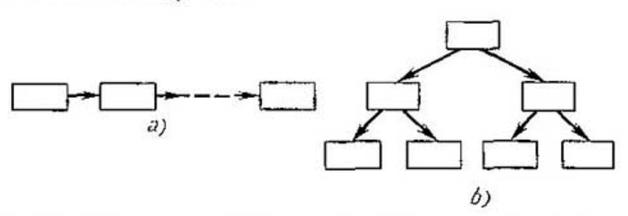


Fig. I-6. Sistemas con enlaces sucesivos (*a*) y estructura jerárquica

¿Cómo formularía Ud. la regla para convertir un número de la forma decimal en la binaria?

I-4. Escriba en los sistemas de numeración binario, óctuple y sexadecimal los números siguientes (dados en el sistema de numeración decimal):

27, 467, 519, 1 263.

Nota. Emplear las letras *a*, *b*, *c*, *d*, *e*, *f* del alfabeto romano como cifras del 10 al 15 del sistema de numeración sexadecimal.

I-5. Escriba los números del problema I-4 en los sistemas de numeración decimal-binario y óctuple-binario. ¿Cuántas cifras binarias tuvo que utilizar Ud. para escribir una cifra óctuple? ¿Cuál de estos dos sistemas de notación de números es más económico desde el punto de vista del número de cifras binarias utilizadas?

I-6. Tienen o no circuito de realimentación los siguientes sistemas de control del tráfico urbano:

1) con ayuda de un semáforo en el cual se encienden sucesivamente las luces roja, amarilla y verde durante un tiempo previamente convenido;

2) con ayuda de un policía de tráfico.

I-7. Halle el circuito de realimentación en los esquemas de las figuras I-4 e I-5.

INDICE DE NOTACIONES

$\{\ldots\}$	conjunto
$\in$	pertenencia al conjunto
$\notin$	no pertenencia al conjunto
$\subseteq$	símbolo de inclusión
$\subset$	símbolo de inclusión estricta
$\cup$	unión de conjuntos
$\cap$	intersección de conjuntos
$\setminus$	diferencia de conjuntos
$\times$	producto directo de conjuntos
$\overline{X}$	complemento del conjunto X
X^s	potencia del conjunto
$\varnothing$	conjunto vacío
I	conjunto universal
R	conjunto de números reales
$\mathfrak{M}$	sistema de conjuntos
$\sup M$	límite superior del conjunto M
$\inf M$	límite inferior del conjunto M
$\mathrm{Pr}_i M$	proyección del conjunto sobre el eje i
$(a_1, \ldots, a_n)$	conjunto ordenado (cortejo, vector)
Λ	cortejo vacío
$\forall$	cuantificador de generalidad
$\rightleftarrows$	equivalencia por definición
$\rightarrow$	consecuencia, reflejo
$f \circ g$	composición de las funciones f y g
$\equiv$	símbolo de relación de equivalencia
$\preccurlyeq$	símbolo de relación de orden
$\prec$	símbolo de relación de orden estricto
$\gg$	símbolo de relación de predominio
$d(x, y)$	distancia entre elementos del conjunto
$\|x\|$	norma de la magnitud x
E_n	espacio euclídeo n-dimensional
$C_{[a, b]}$	espacio de funciones continuas
x/y	función de Sheffer
$\bar{x}$	negación de la enunciación x
$\oplus$	adición por el módulo 2

Z	espacio de resultados del experimento
z	resultado del experimento, magnitud aleatoria
$\mathring{z}$	magnitud aleatoria centrada
$p(z),\ z \in Z$	distribución de las probabilidades en el espacio Z
$P(S),\ P_S$	probabilidad de un suceso
$p(z/S)$	distribución condicional de probabilidades
$p(y, z)$	distribución conjunta de probabilidades
$\bar{z},\ \nu,\ M(z),\ \overline{f(z)}$	valor medio
$\sigma,\ \sigma_z$	desviación cuadratica media
$R(\nu, \omega)$	distribución uniforme
$N(\nu, \sigma^2)$	distribución normal
$w(r, n, p)$	distribucion binomial
$w(r, a)$	distribución de Poisson
u	control
ϑ	estado de la naturaleza
x	estado del sistema
$T(x, u)$	transformación de estado del sistema
$q(x, u),\ Q(x, u)$	función de objetivo
$J(u),\ J_n(u)$	criterio de calidad de control
$\xi,\ \eta$	estrategias mixtas
$L(\xi, \eta)$	función de pérdidas
A	espacio de decisiones
a	decisión, operación
a^*	operación de Bayes
$R^*(\xi)$	pérdidas con la operación de Bayes
$d(z)$	función de decisión
$\rho(\vartheta, d)$	función de riesgo
$\rho^*(\xi)$	riesgo de Bayes

Parte primera

FUNDAMENTOS DE MATEMÁTICAS DISCRETAS

Capítulo primero

CONCEPTOS BASICOS DE LA TEORIA DE LOS CONJUNTOS

1-1. CONJUNTOS FINITOS E INFINITOS

a) Definiciones fundamentales

El concepto de conjunto es uno de aquellos conceptos fundamentales de las matemáticas a los cuales es difícil dar una definición precisa utilizando conceptos elementales. Por eso nos limitaremos a dar una explicación descriptiva del concepto de conjunto. Se denomina *conjunto* a un grupo de objetos determinados completamente diferenciables que se consideran como un todo.

Se puede hablar del conjunto de sillas en una habitación, el conjunto de personas que viven en la ciudad de Riazán, el conjunto de estudiantes en un grupo, el conjunto de números enteros positivos, el conjunto de letras en el alfabeto, el conjunto de estados de un sistema, etc. Por cierto, sólo se puede hablar de conjunto cuando los elementos del mismo son diferenciables entre sí. Por ejemplo, no se puede hablar del conjunto de gotas en un vaso de agua ya que no es posible señalar, precisa y claramente, cada gota por separado.

Llámanse elementos los distintos objetos de que consta un conjunto. Así, el número 3 es un elemento del conjunto de los números enteros positivos y la letra *б* es un elemento del conjunto de las letras del alfabeto ruso.

Sirve de notación general de conjunto, un par de corchetes { }, dentro de los cuales se enumeran los elementos del conjunto. Para designar los conjuntos concretos se utilizan distintas letras mayúsculas A, S, X ... o letras mayúsculas con subíndices A_1, A_2 ...

Para designar los elementos de un conjunto en forma general se emplean diversas letras minúsculas a, s, x ... o letras minúsculas con subíndices a_1, a_2 ...

Si deseamos indicar que el objeto a es un elemento del conjunto S escribimos $a \in S$ y decimos que el objeto a es un elemento del conjunto S o que a pertenece al conjunto S. Ahora, si queremos señalar que a no es un elemento del conjunto S, entonces escribimos $a \notin S$. Vamos a emplear la notación $x_1, x_2, \ldots, x_n \in S$ en calidad de abreviatura para la notación $x_1 \in S, x_2 \in S, \ldots, x_n \in S$.

Los conjuntos pueden ser finitos e infinitos. Un conjunto se denomina *finito* si es finito el número de sus elementos, es decir, si existe un número entero positivo N que es el número de elementos del conjunto. Un conjunto se llama *infinito*, si consta de un número infinito de elementos.

Para operar con conjuntos concretos hay que saber presentar dichos conjuntos. Existen dos métodos de presentación de conjuntos: la enumeración y la descripción. La presentación de un conjunto por el método de enumeración consiste en la enumeración de todos los elementos que constituyen el conjunto. Así, el conjunto de estudiantes sobresalientes de un grupo estudiantil puede presentarse enumerando los estudiantes que obtienen notas de sobresaliente, por ejemplo {Ivanov, Petrov, Sídorov}. Este método es conveniente para el estudio de los conjuntos finitos que contienen un número reducido de elementos, pero a veces se puede utilizar para presentar conjuntos infinitos, por ejemplo $\{2, 4, 6, 8 \ldots\}$. Naturalmente, tal notación es aplicable si está del todo claro qué se comprende por los puntos suspensivos.

El método descriptivo de presentación de un conjunto consiste en que se indica la propiedad característica que poseen todos los elementos del conjunto. Así, si M es el conjunto de estudiantes de un grupo, el conjunto A de estudiantes sobresalientes de dicho grupo se escribe en la forma

$$A = \{x \in M : x, \text{ estudiante sobresaliente del grupo}\},$$

lo que se lee del modo siguiente: el conjunto A consta de los elementos x del conjunto M que poseen la propiedad de que x es estudiante sobresaliente del grupo.

En los casos en que no suscita dudas de cuál conjunto se toman los elementos x, puede no indicarse la pertenencia de x al conjunto M. En tal caso el conjunto A se escribirá en la forma

$$A = \{x : x, \text{ estudiante sobresaliente del grupo}\}.$$

Pongamos algunos ejemplos de presentación de conjuntos por el método descriptivo:

$\{x : x, \text{ par}\}$ es el conjunto de los números pares;

$\{x : x^2 - 1 = 0\}$ es el conjunto $\{+1, -1\}$.

Supongamos que C es el conjunto de los números enteros. Entonces $\{x \in C : 0 < x \leqslant 7\}$ es el conjunto $\{1, 2, 3, 4, 5, 6, 7\}$.

El concepto de conjunto vacío es un importante concepto de la teoría de los conjuntos. Se denomina *conjunto vacío* el que no contiene ningún elemento. El conjunto vacío se designa por $\varnothing$. Por ejemplo,

$$\{x \in C : x^2 - x + 1 = 0\} = \varnothing.$$

El concepto de conjunto vacío juega un papel importantísimo al presentar un conjunto por medio de su descripción. Sin el concepto de conjunto vacío no podríamos hablar del conjunto de estudiantes sobresalientes de un grupo estudiantil o de raíces reales de una ecuación cuadrática sin habernos cerciorado previamente de si hay estudiantes sobresalientes algunos en el grupo dado o de si la ecuación considerada tiene raíces reales. La introducción del conjunto vacío permite operar con plena tranquilidad con el conjunto de estudiantes sobresalientes del grupo sin preocuparse de si hay o no estudiantes sobresalientes en el grupo examinado. Vamos a incluir convencionalmente los conjuntos vacíos en los conjuntos finitos.

Consideremos ahora la cuestión de la igualdad de los conjuntos. Dos conjuntos se denominan *iguales* si constan de los mismos elementos, o sea, si vienen a ser un mismo conjunto. Los conjuntos X e Y no son iguales ($X \neq Y$), si en el conjunto X hay elementos que no pertenecen a Y, o en el conjunto Y existen elementos no pertenecientes a X. Es fácil ver que para conjuntos cualesquiera X, Y y Z:

$$X = X;$$

$$\text{si } X = Y, \text{ entonces } Y = X;$$

$$\text{si } X = Y \text{ e } Y = Z, \text{ entonces } X = Z.$$

De la definición de igualdad de conjuntos se desprende que no importa el orden de los elementos en el conjunto. Así, por ejemplo, los conjuntos $\{3, 4, 5, 6\}$ y $\{4, 5, 6, 3\}$ resultan ser un mismo conjunto.

Con frecuencia, al estudiar diversos conjuntos es preciso referirse al número de elementos del conjunto. Para que este concepto quede definido por completo es necesario convenir en que en un conjunto nunca hay elementos iguales. La notación $\{2, 2, 3, 5\}$ debe considerarse incorrecta y sustituirse por $\{2, 3, 5\}$. De este modo, el conjunto de los divisores primos del número 60 es igual a $\{2, 3, 5\}$.

b) Concepto de subconjunto

El conjunto X constituye un subconjunto del conjunto Y, si cualquier elemento del conjunto X pertenece también al conjunto Y. Sea Y el conjunto de los estudiantes de un grupo y X, el conjunto de los estudiantes sobresalientes del mismo grupo. Como que cada

estudiante sobresaliente del grupo es a la vez estudiante de dicho grupo, el conjunto X es un subconjunto del conjunto Y.

Muchas definiciones de la teoría de los conjuntos se exponen cómodamente en forma de expresiones matemáticas que contienen algunos símbolos lógicos. Utilizaremos dos de estos símbolos para definir el subconjunto:

$\forall$, símbolo denominado cuantificador y que significa "cualquier", "cualquiera que sea", "para todos";

$\rightarrow$, símbolo de corolario (implicación) que significa "implicar".

La definición de subconjunto que puede formularse así: para cualquier x la afirmación «x pertenece a X" implica la aseveración "x pertenece a Y", se escribirá del modo siguiente:

$$\forall x: x \in X \rightarrow x \in Y. \tag{1-1}$$

Una notación más concisa de la expresión "X es un subconjunto de Y" es la transcripción

$$X \subseteq Y, \tag{1-2}$$

que se lee "Y incluye a X". El símbolo $\subseteq$ empleado aquí denota inclusión. Si se desea recalcar que Y incluye también a otros elementos además de los elementos de X, se utiliza el símbolo de inclusión estricta $\subset$:

$$X \subset Y. \tag{1-3}$$

La relación entre los símbolos $\subset$ y $\subseteq$ se muestra con la expresión

$$X \subset Y \rightleftarrows X \subseteq Y \text{ y } X \neq Y. \tag{1-4}$$

Aquí se emplea el símbolo $\rightleftarrows$ que denota equivalencia (con el significado "es lo mismo que").

Mencionemos algunas propiedades de los subconjuntos que se derivan de su definición:

$$X \subseteq X \qquad \text{(reflexividad);}$$

$$[X \subseteq Y \text{ e } Y \subseteq Z] \rightarrow X \subseteq Z \qquad \text{(transitividad).}$$

Es algo más difícil de comprender que para cualquier conjunto M

$$\varnothing \subseteq M. \tag{1-5}$$

En efecto, el conjunto vacío $\varnothing$ no contiene elementos. Por consiguiente, al adicionar a M un conjunto vacío, de hecho no añadimos nada. Por eso, siempre se puede considerar que cualquier conjunto M contiene en sí un conjunto vacío en calidad de subconjunto.

c) Límites superior e inferior de un conjunto

Al operar con el conjunto de los números reales pueden compararse en magnitud los elementos de dicho conjunto. Al hacerlo, surge con frecuencia el problema de determinar el elemento mayor

o menor del conjunto. Para un conjunto finito este problema no presenta dificultad. Así, para el conjunto $T = \{3, 4, 5, 6\}$ mín $T = 3$ y máx $T = 6$. Sin embargo, para los conjuntos infinitos la situación puede ser diferente.

Sea R el conjunto de todos los números reales, siendo $S = \{x \in R : m < x < M\}$. El conjunto S constituye un segmento abierto del eje real y no contiene los elementos mayor y menor. No obstante, puede hablarse de los límites de tal conjunto entendiéndose por éstos los números m y M que complementan el conjunto S hasta el segmento cerrado. En tal caso el punto M se denomina límite superior del conjunto S o *supremum* y se designa por $M = \sup S$, y el punto m se llama límite inferior del conjunto S o *infinum* y se designa con $m = \inf S$.

Teorema 1-1 (teorema de los límites superior e inferior de un subconjunto). Si $B \subseteq A$, entonces

$$\inf B \geqslant \inf A; \quad \sup B \leqslant \sup A. \tag{1-6}$$

Demostración. Designemos por b' el elemento del conjunto B que tiene el valor menor, o sea, $b' \in B$ y $b' = \inf B$. Pero $B \subseteq A$, es decir, $b' \in A$. Sea a' el elemento del conjunto A que posee el menor valor, o sea, $a' \in A$ y $a' = \inf A$. En tal caso, si $b' = a'$, entonces $b' = \inf A$ y si $b' \neq a'$, entonces $b' > a = \inf A$. De este modo, $b' \geqslant \inf A$ o $\inf B \geqslant \inf A$.

La segunda parte del teorema se demuestra análogamente.

1-2. OPERACIONES CON CONJUNTOS

a) Observaciones preliminares

Con conjuntos pueden realizarse operaciones que recuerdan en mucho las operaciones de adición y multiplicación del álgebra elemental. Para comprender mejor las operaciones con conjuntos hay que rememorar las leyes existentes en el álgebra elemental.

Sean a y b ciertos números, $a + b$ su suma y ab su producto. La suma y el producto de estos números tienen las propiedades siguientes denominadas leyes del álgebra:

1. $a + b = b + a$; $ab = ba$, ley conmutativa;
2. $(a + b) + c = a + (b + c)$; $(ab)c = a(bc)$, ley asociativa;
3. $(a + b)c = ac + bc$, ley distributiva.

Observemos que en las leyes asociativa y conmutativa la operación de adición puede sustituirse por la de multiplicación y viceversa. En tal caso se obtendrá otra ley que será tan válida como la primera. Sin embargo, en la ley distributiva no existe tal simetría. Si en dicha ley se reemplaza la adición por la multiplicación y viceversa, llegamos al absurdo:

$$(ab) + c = (a + c)(b + c).$$

Cabe preguntar, ¿ acaso sucede siempre así? ¿No existe acaso un álgebra en la cual la ley distributiva sea también simétrica con respecto a las operaciones de adición y multiplicación como las leyes conmutativa y asociativa? Resulta que existe un álgebra, y precisamente el álgebra de los conjuntos, en la que todas las tres leyes son simétricas respecto a las operaciones de adición y multiplicación.

La similitud entre las operaciones de adición y multiplicación se manifiesta asimismo, en la existencia de dos excelentes números que son 0 y 1 tales que la adición del primero y la multiplicación por el segundo no modifican ningun número:

$$a+0=a, \quad a\cdot 1=a.$$

Notemos que la segunda relación se obtiene de la primera sustituyendo $(+)$ por $(\cdot)$ y 0 por 1.

No obstante, en este asunto tampoco va muy lejos el parecido entre las operaciones de adición y multiplicación. Así, el número 0 juega un papel algo particular en comparación con todos los demás números, incluso la unidad. Este papel peculiar del número 0 se deriva de la relación $a\cdot 0=0$. Si en dicha expresión sustituimos $(\cdot)$ por $(+)$ y 0 por 1 llegamos a la relación $a+1=1$ que casi nunca será correcta.

Como veremos más adelante, la similitud entre el cero y la unidad resultará mucho mayor en el álgebra de los conjuntos que en el álgebra ordinaria.

Después de estas observaciones preliminares puede iniciarse el estudio de las operaciones con conjuntos.

b) Unión de conjuntos

Se denomina unión de los conjuntos X e Y aquel que consta de todos aquellos y sólo aquellos elementos que pertenecen siquiera a uno de los conjuntos X, Y, es decir, pertenecen a X o a Y. La unión de X e Y se designa por $X\cup Y$. La definición formal es

$$X\cup Y=\{x: x\in X \text{ o } x\in Y\}. \tag{1-7}$$

La unión de conjuntos se llama a veces suma de conjuntos y se designa por $X+Y$. No obstante, las propiedades de la unión de conjuntos difieren un tanto de las propiedades de la suma en su sentido aritmético habitual. Por eso no vamos a emplear este último término.

Ejemplo 1-1. Si $X=\{1, 2, 3, 4, 5\}$ e $Y=\{2, 4, 6, 7\}$, entonces $X\cup Y=\{1, 2, 3, 4, 5, 6, 7\}$.

Ejemplo 1-2. Si X es el conjunto de estudiantes sobresalientes en un grupo, e Y, el conjunto de estudiantes que viven en una residencia estudiantil, entonces $X\cup Y$ es el conjunto de estudiantes que cursan estudios con notas de sobresaliente o habitan en dicha residencia.

Ejemplo 1-3. Examinemos los dos círculos mostrados en la figura 1-1. Si X es el conjunto de puntos de círculo izquierdo e Y, el conjunto de puntos del círculo derecho, entonces $X \cup Y$ constituye la zona rayada delimitada por ambos círculos.

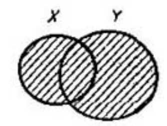


Fig. 1-1. Unión de conjuntos

El concepto de unión puede hacerse extensivo también a un número mayor de conjuntos. Designemos por $\mathfrak{M} = \{X_1, \ldots, X_n\}$ la totalidad de los n conjuntos $X_1, \ldots, X_n$ que a veces se denomina sistema de conjuntos. La unión de estos conjuntos

$$\bigcup_{i=1}^{n} X_i = \bigcup_{X \in \mathfrak{M}}, \quad X = X_1 \cup \ldots \cup X_n \tag{1-8}$$

representa el conjunto que consta de todos aquellos y sólo aquellos elementos que pertenecen siquiera a uno de los conjuntos del sistema $\mathfrak{M}$.

Para la unión de conjuntos son válidas las leyes conmutativa y asociativa

$$X \cup Y = Y \cup X; \tag{1-9}$$

$$(X \cup Y) \cup Z = X \cup (Y \cup Z) = X \cup Y \cup Z, \tag{1-10}$$

cuya validez se deriva de que los miembros izquierdo y derecho de las igualdades constan de los mismos elementos. Además,

$$X \cup \varnothing = X. \tag{1-11}$$

Esta relación es también evidente ya que un conjunto vacío no contiene elementos y esto significa que X y $X \cup \varnothing$ constan de los mismos elementos. En (1-11) se ve que el conjunto vacío $\varnothing$ juega en el álgebra de los conjuntos el papel de cero. Aquí existe analogía con la expresión $a + 0 = a$ del álgebra ordinaria.

c) Intersección de conjuntos

Se denomina intersección de los conjuntos X e Y el conjunto que consta de todos aquellos y sólo aquellos elementos que pertenecen tanto al conjunto X como al Y. La intersección de los conjuntos X e Y se designa por $X \cap Y$. La definición formal es

$$X \cap Y \{x : x \in X \text{ y } x \in Y\}. \tag{1-12}$$

La intersección de conjuntos se denomina a veces producto de conjuntos y se designa por XY. Sin embargo, las propiedades de

la intersección de conjuntos difieren un tanto de las propiedades del producto en su sentido aritmético habitual. Por eso no vamos a emplear dicho término.

Ejemplo 1-4. Para los conjuntos X e Y en el ejemplo 1-1 $X \cap Y = \{2, 4\}$.

Ejemplo 1-5. Para los conjuntos X e Y en el ejemplo 1-2 $X \cap Y$ es el conjunto de estudiantes sobresalientes que viven en la residencia estudiantil.

Ejemplo 1-6. Examinemos los dos círculos mostrados en la figura 1-2. Si X es el conjunto de puntos de círculo izquierdo e Y, el conjunto de puntos del

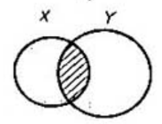


Fig. 1-2. Intersección de conjuntos

círculo derecho, entoces $X \cap Y$ representa la zona rayada que constituye la parte común de ambos círculos.

La operación de intersección permite establecer una serie de relaciones entre dos conjuntos.

Los conjuntos X e Y se denominan *no intersecados* si éstos no tienen elementos comunes, es decir, si

$$X \cap Y = \varnothing. \qquad (1\text{-}13)$$

Ejemplo 1-7. Son conjuntos no intersecados:

1) los conjuntos $\{1, 2, 3\}$ y $\{4, 5, 6\}$;

2) el conjunto de estudiantes sobresalientes y el conjunto de estudiantes desaprovechados en un grupo estudiantil;

3) los conjuntos de puntos de los círculos X e Y en la figura 1-3.

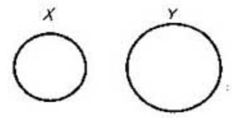


Fig. 1-3. Conjuntos no intersecados

Se dice que los conjuntos X e Y se encuentran *en posición común* si se cumplen las tres condiciones siguientes:

existe un elemento del conjunto X que no pertenece a Y;

existe un elemento del conjunto Y que no pertenece a X;

existe un elemento que pertenece tanto a X como a Y.

Señalemos una diferencia entre el álgebra de los conjuntos y el álgebra de los números. Si a y b son dos números, entre ellos pueden existir tres relaciones o posibilidades

$$a < b, \quad a = b, \quad b < a. \qquad (1\text{-}14)$$

Sin embargo, para dos conjuntos X e Y puede no cumplirse ninguna de las relaciones:

$$X \subset Y, \quad X = Y, \quad Y \subset X. \tag{1-15}$$

Así pues, si X es el conjunto de estudiantes sobresalientes e Y es el conjunto de estudiantes que viven en la residencia estudiantil, las tres relaciones antes dadas significan:

$X \subset Y$, cada estudiante sobresaliente vive obligatoriamente en la residencia estudiantil;

$X = Y$, en la residencia estudiantil viven todos los estudiantes sobresalientes y sólo ellos;

$Y \subset X$, todos los estudiantes que habitan en la residencia estudiantil son estudiantes sobresalientes.

Es evidente que estas relaciones no agotan todas las posibilidades. En efecto, como se desprende de las definiciones anteriores, entre los dos conjuntos X e Y puede haber una de las relaciones siguientes:

$$X = Y; \quad X \subset Y; \quad Y \subset X, \quad X \cap Y = \varnothing;$$

X e Y están en posición común.

Se puede extender también el concepto de intersección a un número de conjuntos mayor de dos. Examinemos el sistema de conjuntos $\mathfrak{M} = \{X_1, \ldots, X_n\}$. La intersección de estos conjuntos se escribe en la forma

$$\bigcap_{X \in \mathfrak{M}} X = \bigcup_{i=1}^{n} X_i = X_1 \cap \ldots \cap X_n \tag{1-16}$$

y representa un conjunto cuyos elementos pertenecen a cada uno de los conjuntos del sistema $\mathfrak{M}$.

Es fácil ver que la intersección de conjuntos posee la propiedad conmutativa

$$X \cap Y = Y \cap X \tag{1-17}$$

y la asociativa

$$(X \cap Y) \cap Z = X \cap (Y \cap Z) = X \cap Y \cap Z. \tag{1-18}$$

Notemos también que tiene lugar la relación

$$X \cap \varnothing = \varnothing, \tag{1-19}$$

similar a la relación $a \cdot 0 = 0$ del álgebra ordinaria. La relación (1-19), junto con la relación (1-11), demuestra que el conjunto vacío juega el papel de cero en el álgebra de los conjuntos.

d) Diferencia de conjuntos

La operación dada se distingue de las operaciones de unión e intersección en que se define solamente para dos conjuntos. Se denomina diferencia de los conjuntos X e Y el conjunto que consta de todos aquellos y sólo aquellos elementos que pertenecen a X y

no pertenecen a Y. La diferencia de los conjuntos X e Y se designa por $X \setminus Y$. De este modo,

$$X \setminus Y = \{x : x \in X, \quad x \notin Y\}. \tag{1-20}$$

Ejemplo 1-8. Para los conjuntos X e Y del ejemplo 1-1 $X \setminus Y = \{1, 3, 5\}$, $Y \setminus X = \{6, 7\}$. Si X e Y son los conjuntos del ejemplo 1-2, entonces $X \setminus Y$ es el conjunto de estudiantes sobresalientes que no viven en la residencia estudiantil. Para los conjuntos X e Y del ejemplo 1-3 $X \setminus Y$ es la zona rayada de la figura 1-4.

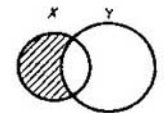


Fig. 1-4. Diferencia de conjuntos

e) Conjunto universal

Como hemos visto, el conjunto vacío representa el papel de cero en el álgebra de los conjuntos. Puede preguntarse si no existe acaso un conjunto I que juegue el papel de unidad, o sea, que satisfaga la condición

$$X \cap I = X, \tag{1-21}$$

similar a la condición $a \cdot 1 = a$ del álgebra ordinaria.

La relación (1-21) significa que la intersección o "parte común" del conjunto I y del conjunto X, para cualquier conjunto X coincide con este mismo conjunto. Pero eso sólo es posible en caso de que el conjunto I contenga todos los elementos de los cuales puede constar el conjunto X, de suerte que cualquier conjunto X esté totalmente contenido en el conjunto I. El conjunto I que satisface esta condición se denomina *completo* o *universal*, o bien *unitario.*

Partiendo de lo expuesto puede darse la siguiente definición de conjunto universal. Si en cierto análisis sólo participan los subconjuntos de cierto conjunto invariable I, entonces dicho conjunto mayor I se denomina conjunto universal.

Es preciso señalar que en los diversos análisis concretos pueden representar el papel de conjunto universal diferentes conjuntos. Así, pues, al estudiar los conjuntos de estudiantes en un grupo (estudiantes sobresalientes, estudiantes becados, estudiantes que viven en la residencia estudiantil, etc.), juega el papel de conjunto universal el conjunto de estudiantes en el grupo.

Es cómodo representar gráficamente el conjunto universal en forma del conjunto de los puntos de un rectángulo. Las distintas zonas dentro de dicho rectángulo van a denotar los diversos subconjuntos del conjunto universal. La representación de los conjun-

tos en forma de zonas de un rectángulo que es un conjunto universal se denomina *diagrama de Euler — Venn.*

El conjunto universal posee una interesante propiedad que no tiene análogo en el álgebra ordinaria, a saber: para cualquier conjunto X es válida la relación:

$$X \cup I = I. \qquad (1\text{-}22)$$

Efectivamente, la unión $X \cup I$ constituye el conjunto del cual forman parte todos los elementos tanto del conjunto X como del I. Pero el conjunto I ya incluye todos los elementos del conjunto X, por lo que $X \cup I$ va a constar de los mismos elementos que I, o sea, constituye el conjunto más universal I.

f) Complemento de un conjunto

El conjunto $\bar{X}$ definido por la relación

$$\bar{X} = I \setminus X, \qquad (1\text{-}23)$$

se denomina complemento del conjunto X (hasta el conjunto universal I). En el diagrama de la figura 1-5 la zona no rayada re-

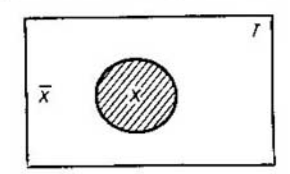


Fig. 1-5. Complemento de un conjunto

presenta el conjunto $\bar{X}$. La definición formal es

$$\bar{X} = \{x : x \in I \text{ y } x \notin X\}.$$

Ejemplo 1-9. Si $I = \{1, 2, 3, 4, 5, 6, 7\}$ y $X = \{3, 5, 7\}$, entonces $\bar{X} = \{1, 2, 4, 6\}$.

De (1-23) se deriva que X y $\bar{X}$ no tienen elementos comunes, de manera que

$$X \cap \bar{X} = \varnothing. \qquad (1\text{-}24)$$

Además, no existen elementos de I que no pertenezcan a X ni a $\bar{X}$, ya que los elementos que no pertenecen a X pertenecen a $\bar{X}$. Por lo tanto,

$$X \cup \bar{X} = I. \qquad (1\text{-}25)$$

De la simetría de la fórmula (1-25) respecto a X y $\bar{X}$ se deduce no sólo que $\bar{X}$ es el complemento de X, sino también que X es el

complemento de $\bar{X}$. Pero el complemento de $\bar{X}$ es $\bar{\bar{X}}$. De este modo,

$$\bar{\bar{X}} = X. \tag{1-26}$$

Con ayuda de la operación de complementación puede representarse la diferencia de conjuntos en la cómoda forma siguiente:

$$X \setminus Y = \{x : x \in X \text{ y } x \notin Y\} = \{x : x \in X \text{ y } X \in \bar{Y}\},$$

o sea,

$$X \setminus Y = X \cap \bar{Y}. \tag{1-27}$$

g) Fraccionamiento de un conjunto

Una de las operaciones con conjuntos que se encuentran más frecuentemente es la operación de fraccionamiento de un conjunto en un sistema de subconjuntos. Así pues, el sistema de cursos de una facultad dada constituye el fraccionamiento del conjunto de estudiantes de dicha facultad; el sistema de los grupos de un curso dado es el fraccionamiento del conjunto de estudiantes del curso. Si N es el conjunto de los números enteros positivos, siendo A_0 y A_1 los conjuntos de los números pares e impares, entonces el sistema $\{A_0, A_1\}$ será el fraccionamiento del conjunto N. El conjunto de los números enteros positivos puede fraccionarse de otro modo, a saber: en los conjuntos de los números divisibles entre 3 sin resto, con resto 1 y con resto 2. La producción de una empresa se fracciona en el sistema de conjuntos que constan de los productos de primera clase, segunda clase y artículos defectuosos. Se podrían citar infinitos ejemplos similares.

Para dar una definición regurosa del concepto de fraccionamiento, examinemos cierto conjunto M y el sistema de conjuntos $\mathfrak{M} = \{X_1, \ldots, X_n\}$. El sistema de conjuntos $\mathfrak{M}$ se denomina fraccionamiento del conjunto M si satisface las condiciones siguientes:

1) cualquier conjunto X de $\mathfrak{M}$ es subconjunto del conjunto M

$$\forall X \in \mathfrak{M} : X \subseteq M; \tag{1-28}$$

2) dos conjuntos cualesquiera X e Y de $\mathfrak{M}$ no son intersecados

$$\forall X \in \mathfrak{M}, \quad \forall Y \in \mathfrak{M} : X \neq Y \to X \cap Y = \varnothing; \tag{1-29}$$

3) la unión de todos los conjuntos que forman parte del fraccionamiento da el conjunto M

$$\bigcup_{X \in \mathfrak{M}} X = M. \tag{1-30}$$

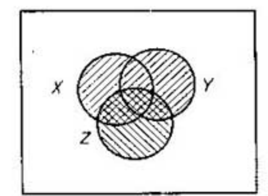
Volveremos al concepto de fraccionamiento durante el estudio de la relación de equivalencia, con la que aquél está estrechamente vinculado.

h) Identidades del álgebra de los conjuntos

Con ayuda de las operaciones de unión, intersección y complementación es posible formar diversas expresiones algebraicas de los conjuntos. Designemos por $\mathfrak{A}(X, Y, Z)$ cierta expresión algebraica constituida por los conjuntos X, Y y Z. Esta representa cierto conjunto. Sea $\mathfrak{B}(X, Y, Z)$ otra expresión algebraica formada por los mismos conjuntos. Si ambas expresiones algebraicas representan un mismo conjunto, entonces se les puede igualar una a otra obteniendo una identidad algebraica de la forma

$$\mathfrak{A}(X, Y, Z) = \mathfrak{B}(X, Y, Z). \qquad (1\text{-}31)$$

Tales conjuntos resultan sumamente útiles para las transformaciones de las expresiones algebraicas con conjuntos y en la sección presente estudiaremos algunos de ellos.



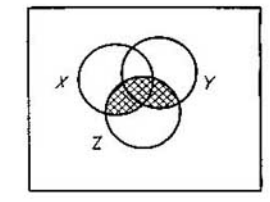


Fig. 1-6. Representación geométrica de la identidad $(X \cup Y) \cap Z = (X \cap Z) \cup (Y \cap Z)$

1. En la figura 1-6 se muestran los diagramas de Euler — Venn para las expresiones $(X \cup Y) \cap Z$ y $(X \cap Z) \cup (Y \cap Z)$. En dichos diagramas se ve que ambas expresiones definen un mismo conjunto, así pues, en el álgebra de los conjuntos tiene lugar la identidad

$$(X \cup Y) \cap Z = (X \cap Z) \cup (Y \cap Z), \qquad (1\text{-}32)$$

análoga a la ley distributiva $(a + b)c = ac + bc$ del álgebra ordinaria.

2. En el álgebra ordinaria, en la ley distributiva no podemos sustituir la operación de adición por la de multiplicación, y viceversa, ya que esto conduce a la expresión absurda $(ab) + c = (a + c)(b + c)$. No ocurre así en el álgebra de los conjuntos.

En la figura 1-7 se muestran los diagramas de Euler — Venn para las expresiones algebraicas $(X \cap Y) \cup Z$ y $(X \cup Z) \cap (Y \cup Z)$. Ambas estas expresiones dan un mismo conjunto; por lo tanto, tiene lugar la identidad

$$(X \cap Y) \cup Z = (X \cup Z) \cap (Y \cup Z). \qquad (1\text{-}33)$$

3. Es fácil cerciorarse de que si $Y \subseteq X$, entonces

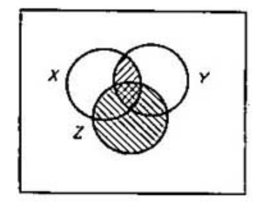
$$X \cap Y = Y, \quad X \cup Y = X. \tag{1-34}$$

En efecto, todos los elementos del conjunto Y son también a la vez, elementos del conjunto X. Esto significa que la intersección de dichos conjuntos, o sea, la parte común de los conjuntos X e Y coincide con Y. En la unión de los conjuntos X e Y, el conjunto Y no aporta ningún elemento que ya no forme parte de la misma siendo elemento del conjunto X. Por consiguiente, $X \cup Y$ coincide con X.

4. Suponiendo que en (1-34) $Y = X$ y considerando que $X \subseteq X$, hallamos que:

$$X \cap X = X, \quad X \cup X = X. \tag{1-35}$$

En algunos casos resulta inconveniente demostrar las identidades del álgebra de los conjuntos con ayuda del diagrama de



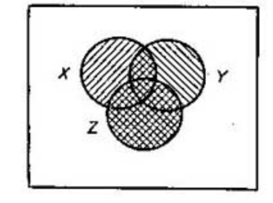


Fig. 1-7. Representación geométrica de la identidad $(X \cap Y) \cup Z = (X \cup Z) \cap (Y \cup Z)$

Euler — Venn. Existe un método más general para determinar la identidad de dos expresiones algebraicas.

Sean designadas, como antes, por $\mathfrak{A}(X, Y, Z)$ y $\mathfrak{B}(X, Y, Z)$, dos expresiones algebraicas obtenidas mediante la aplicación de las operaciones de unión, intersección y complementación a los conjuntos. X, Y y Z. Para demostrar que $\mathfrak{A} = \mathfrak{B}$ es suficiente mostrar que $\mathfrak{A} \subseteq \mathfrak{B}$ y que $\mathfrak{B} \subseteq \mathfrak{A}$. A su vez, para mostrar que $\mathfrak{A} \subseteq \mathfrak{B}$, es necesario convencerse que de $x \in \mathfrak{A}$ se desprende que $x \in \mathfrak{B}$. Análogamente, para mostrar que $\mathfrak{B} \subseteq \mathfrak{A}$ hace falta cerciorarse de que de $x \in \mathfrak{B}$ se deriva que $x \in \mathfrak{A}$. Utilicemos este método para demostrar varias identidades más.

5. Demostremos la identidad

$$\overline{X \cup Y} = \bar{X} \cap \bar{Y}. \tag{1-36}$$

Supongamos que $x \in \overline{X \cup Y}$, o sea, que $x \notin X \cup Y$. Esto significa que $x \notin X$ y $x \notin Y$, es decir, que $x \in \bar{X}$ y $x \in \bar{Y}$. Luego, $x \in \bar{X} \cap \bar{Y}$. Supongamos ahora que $y \in \bar{X} \cap \bar{Y}$, o sea, $y \in \bar{X}$ e $y \in \bar{Y}$. Esto quiere decir que $y \notin X$ e $y \notin Y$, o bien que $y \notin X \cup Y$. Por lo tanto, $y \in \overline{X \cup Y}$.

6. Demostremos la identidad

$$\overline{X \cap Y} = \bar{X} \cup \bar{Y}, \qquad (1\text{-}37)$$

reduciendo ambos miembros de la misma a igual forma. Efectuando la operación de complementación en ambos miembros de (1-37) obtenemos que $\overline{\overline{X \cap Y}} = \overline{\bar{X} \cup \bar{Y}}$. El primer miembro de esta expresión da $X \cap Y$. Lo mismo obtenemos transformando el segundo miembro según la regla (1-36).

En la literatura las identidades (1-36) y (1-37) comúnmente se denominan identidades de Morgan.

1-3. ORDENACION DE ELEMENTOS Y PRODUCTO DIRECTO DE LOS CONJUNTOS

a) Conjunto ordenado

A la par con el concepto de conjunto, como una unión de elementos, existe una importante noción de conjunto ordenado o cortejo. Se denomina *cortejo* una sucesión de elementos, o sea, una union de elementos en la cual cada uno de ellos ocupa un lugar determinado. Los propios elementos se llaman entonces componentes del cortejo (primera componente, segunda componente, etc.). Son ejemplos de cortejos: el conjunto de personas que están en una cola, el conjunto de palabras en una frase, los números que expresan la longitud y latitud de un punto en el terreno, etc. En todos estos conjuntos la ubicación de cada elemento está absolutamente determinada y no puede cambiarse arbitrariamente.

El número de elementos del cortejo se denomina su largo. Para designar un cortejo vamos a emplear los paréntesis. Así pues, el conjunto

$$a = (a_1, a_2, \ldots, a_n) \qquad (1\text{-}38)$$

es el cortejo de largo n con los elementos $a_1, \ldots, a_n$. Los cortejos de largo 2 se llaman pares o pares ordenados, los cortejos de largo 3, tríadas, los de largo 4, cuartas, etc. En el caso general, los cortejos de largo n se llaman de orden n o n-arios. Constituyen casos particulares de cortejos el cortejo (a) de largo 1 y el cortejo vacío de largo 0 designado por ( ) o Λ. A diferencia del conjunto ordinario, en el cortejo pueden haber elementos iguales, p. ej.: dos palabras iguales en una frase, valores numéricos iguales de la longitud y la latitud de un punto en el terreno, etc.

En lo sucesivo vamos a estudiar conjuntos ordenados cuyos elementos son números reales. Tales conjuntos ordenados se denominan puntos del espacio o vectores. De este modo, el cortejo (a_1, a_2) puede considerarse como un punto en el plano o un vector trazado desde el origen de coordenadas hasta el punto en cuestión (fig. 1-8, a). Las componentes a_1 y a_2 serán las proyecciones del

vector sobre los ejes *1* y *2*

$$\mathrm{Pr}_1(a_1, a_2) = a_1; \quad \mathrm{Pr}_2(a_1, a_2) = a_2.$$

El cortejo (a_1, a_2, a_3) se puede considerar como un punto en el espacio tridimensional o como un vector espacial trazado del origen de coordenadas a dicho punto (fig. 1-8, *b*). Las proyecciones del vector sobre los ejes de coordenadas son

$$\mathrm{Pr}_i(a_1, a_2, a_3) = a_i, \quad i = 1, 2, 3.$$

Sin embargo, en el caso dado puede hablarse de la proyección simultánea del cortejo sobre dos ejes, por ejemplo *1* y *2*, es decir, sobre el plano de coordenadas. Es fácil ver que dicha proyección constituye un cortejo de dos elementos

$$\mathrm{Pr}_{12}(a_1, a_2, a_3) = (a_1, a_2).$$

Generalizando estos conceptos vamos a considerar el conjunto ordenado de números reales de n elementos $(a_1, \ldots, a_n)$ como un

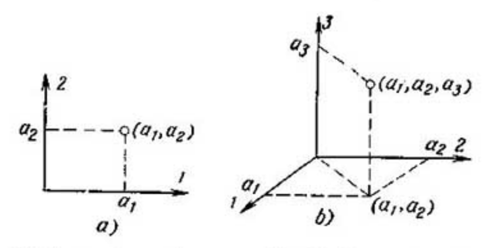


Fig. 1-8. Proyecciones de un cortejo de dos y tres elementos

punto en el espacio imaginario de n dimensiones llamado a veces hiperespacio o como un vector de n dimensiones. Vamos a considerar entonces las componentes del cortejo a de n elementos como las proyecciones de dicho cortejo sobre los ejes correspondientes

$$\mathrm{Pr}_i\, a = a_i, \quad i = 1, \ldots, n. \tag{1.39}$$

Si $i, j, \ldots, l$ son los números de los ejes, siendo $1 \leqslant i < j < \ldots l \leqslant n$, la proyección del cortejo a sobre los ejes $i, j, \ldots, l$ es igual a:

$$\mathrm{Pr}_{i, j, \ldots, l}\, a = (a_i, a_j, \ldots, a_l). \tag{1-40}$$

La proyección del cortejo sobre un conjunto vacío de ejes constituye un cortejo vacío

$$\mathrm{Pr}_{\varnothing}\, a = \Lambda. \tag{1-41}$$

En el capítulo 3 se dará una definición más completa y rigurosa del espacio multidimensional.

b) Producto directo de conjuntos

Se llama producto directo de los conjuntos X e Y al conjunto designado por $X \times Y$ que consta de todos aquellos y sólo aquellos pares ordenados cuya primera componente pertenece al conjunto X y la segunda, al conjunto Y. De esta suerte los elementos del conjunto ordenado constituyen cortejos de dos elementos del tipo (x, y). Su definición formal es

$$X \times Y = \{(x, y) : x \in X,\ y \in Y\}. \tag{1-42}$$

Ejemplo 1-10. Sean $X = \{1, 2\}$, $Y = \{1, 3, 4\}$. Entonces $X \times Y = \{(1, 1), (1,3), (1,4), (2,1), (2,3), (2,4)\}$. La representación geométrica de este conjunto se muestra en la figura 1-9, *a*.

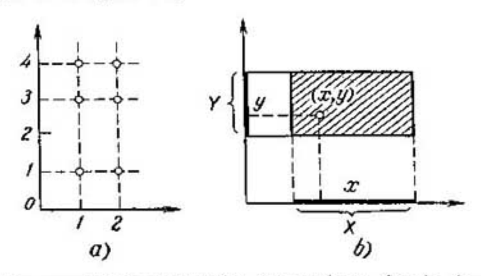


Fig. 1-9. Representación geométrica del producto directo de conjuntos

Ejemplo 1-11. Sean X e Y segmentos de un eje real. El producto directo $X \times Y$ se representa por el rectángulo rayado indicado en la figura 1-9, *b*. De esta figura se deduce que las propiedadaes del producto directo difieren de las propiedades del producto ordinario en sentido aritmético. En particular, el producto directo cambia al cambiar el orden de los factores, es decir,

$$X \times Y \neq Y \times X. \tag{1-43}$$

La operación de multiplicación directa se extiende también fácilmente a un número mayor de conjuntos. Se denomina producto directo de los conjuntos $X_1, X_2, \ldots, X_r$, el conjunto designado por $X_1 \times X_2 \times \ldots X_r$ que consta de todos aquellos y sólo aquellos cortejos de largo r, cuya primera componente pertenece a X_1, la segunda, a X_2, etc.

Es fácil ver que

$$X \times Y = \varnothing \rightleftarrows X = \varnothing \text{ o } Y = \varnothing, \tag{1-44}$$

ya que no existen pares ordenados con la primera o segunda componente que falta. Análogamente, $X_1 \times X_2 \times \ldots X_r = \varnothing$ cuando, y sólo cuando siquiera uno de los conjuntos $X_1, X_2, \ldots, X_r$ es un conjunto vacío.

El concepto de *potencias* de un conjunto constituye un caso particular de la operación de multiplicación directa. Sea M un con-

junto arbitrario. Llamemos potencia s del conjunto M y designemos por M^s el producto directo de s conjuntos iguales a M:

$$M^s = \underbrace{M \times M \times \ldots \times M}_{s \text{ veces}}. \qquad (1\text{-}45)$$

Esta definición sirve para $s = 2, 3, \ldots$ Ella puede hacerse extensiva a cualquier s entero no negativo si asumimos mediante una definición especial que

$$M^1 = M, \quad M^0 = \{\Lambda\}. \qquad (1\text{-}46)$$

Si R es el conjunto de los números reales, entonces $R^2 = R \times R$ representa el plano real y $R^3 = R \times R \times R$ representa el espacio tridimensional real.

c) Proyección de un conjunto

La operación de proyección de un conjunto está vinculada estrechamente a la operación de proyección de un cortejo y sólo puede aplicarse a los conjuntos cuyos elementos son cortejos de largo igual.

Sea M un conjunto que consta de cortejos de largo s. Entonces vamos a llamar proyección del conjunto M al conjunto de las proyecciones de los cortejos de M.

Ejemplo 1-12. Sea $M = \{(1, 2, 3, 4, 5), (2, 1, 3, 5, 5), (3, 3, 3, 3, 3), (3, 2, 3, 4, 3)\}$. Entonces

$$\text{Pr}_2 M = \{2, 1, 3\}; \quad \text{Pr}_{2,4} M = \{(2, 4), (1, 5), (3, 3)\}.$$

Es fácil comprobar que si $M = X \times Y$, entonces

$$\text{Pr}_1 M = X; \quad \text{Pr}_2 M = Y, \qquad (1\text{-}47)$$

y si $Q \subseteq X \times Y$, entonces

$$\text{Pr}_1 Q \subseteq X; \quad \text{Pr}_2 Q \subseteq Y. \qquad (1\text{-}48)$$

1-4. CORRESPONDENCIAS

a) Definición de correspondencia

Examinemos dos conjuntos X e Y. Los elementos de dichos conjuntos pueden compararse unos con otros de algún modo, formando pares (x, y). Si el método de tal comparación está determinado, o sea, para cada elemento $x \in X$ está indicado el elemento $y \in Y$ con el cual se compara el elemento x, se dice que entre los conjuntos X e Y se ha establecido correspondencia no siendo entonces imprescindible en lo absoluto que participen en la comparación todos los elementos de los conjuntos X e Y.

Para presentar una correspondencia es necesario señalar:

1) el conjunto X cuyos elementos se comparan con los elementos del otro conjunto;

2) el conjunto Y cuyos elementos se comparan con los elementos del primer conjunto;

3) el conjunto $Q \subseteq X \times Y$ que define la ley en conformidad con la cual se realiza la correspondencia, o sea, que enumera todos los pares (x, y) participantes en la comparación. De este modo, la correspondencia designada por q representa la tríada de conjuntos

$$q = (X, Y, Q), \tag{1-49}$$

en la cual $Q \subseteq X \times Y$. En esta expresión, la primera componente X se llama dominio de partida de la correspondencia, la segunda componente Y, dominio de llegada de la correspondencia, y la tercera componente Q, gráfica de la correspondencia. El término "gráfica" se explicará más detalladamente al estudiar el tipo particular de correspondencia denominada función.

Además de los tres conjuntos examinados X, Y, Q, también están relacionados indisolublemente con cada correspondencia los dos conjuntos siguientes: el conjunto $\mathrm{Pr}_1 Q$ llamado dominio de definición de la correspondencia, formado por los elementos del conjunto X que entran en comparación y el conjunto $\mathrm{Pr}_2 Q$ llamado dominio de valores de la correspondencia, constituido por los elementos del conjunto Y que entran en comparación.

Si $(x, y) \in Q$, se dice que el elemento y corresponde al elemento x. Es cómodo representar esto geométricamente mediante una flecha dirigida de x a y.

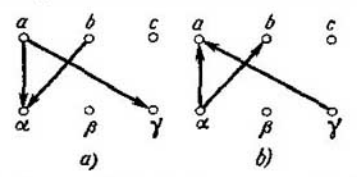


Fig. 1-10. Representación geométrica de las correspondencias directa e inversa

Ejemplo 1-13. Sean $X = \{1, 2\}$, $Y = \{3, 5\}$ de manera que $X \times Y = \{(1, 3), (1, 5), (2, 3), (2, 5)\}$. Este conjunto da la posibilidad de obtener 16 correspondencias diferentes. Citemos algunas de ellas:

$$Q_1 = \{(1, 3)\}; \quad \mathrm{Pr}_1 Q_1 \{1\}; \quad \mathrm{Pr}_2 Q_1 = \{3\};$$

$$Q_2 = \{(1, 3), (1, 5)\}; \quad \mathrm{Pr}_1 Q_2 = \{1\}; \quad \mathrm{Pr}_2 Q_2 = \{3, 5\} = Y.$$

Ejemplo 1-14. En una empresa hay tres vehículos automóviles, dos camiones α y β que trabajan en dos turnos y un autobús γ raramente utilizado. El camión β se encuentra en reparaciones. En la plantilla hay tres chóferes a, b, c de los cuales c está de vacaciones. La distribución de los chóferes por vehículos constituye una correspondencia. Una de las correspondencias posibles será la siguiente:

$$q = (\{a, b, c\}, \{\alpha, \beta, \gamma\}, \{(a, \alpha), (a, \gamma), (b, \alpha)\}).$$

Esta correspondencia está representada geométricamente en la figura 1-10, *a*. En ella el elemento α corresponde a los elementos a y b, y el elemento γ, al elemento a. La correspondncia q está definida para a y b, pero no lo está para c, por consiguiente, el dominio de definición de la correspondencia es el conjunto $\{a, b\}$. El dominio de valores de la correspondencia es el conjunto $\{\alpha, \gamma\}$.

b) Correspondencia inversa

Para cada correspondencia $q=(X, Y, Q)$, $Q \subseteq X \times Y$ existe una correspondencia inversa que se obtiene si se considera la correspondencia dada en dirección contraria, es decir, si se determinan los elementos $x \in X$ con los cuales se comparan los elementos $y \in Y$. La correspondencia inversa a la correspondencia q será designada por

$$q^{-1}=(Y, X, Q^{-1}), \tag{1-50}$$

donde $Q^{-1} \subseteq Y \times X$.

Ejemplo 1-15. Para el ejemplo 1-14 la correspondencia inversa será la distribución de los vehículos entre los chóferes:

$$(\{\alpha, \beta, \gamma\}\ \{a, b, c\},\ \{(\alpha, a), (\alpha, b), (\gamma, a)\}),$$

lo que se muestra geométricamente en la figura 1-10 *b*.

En el ejemplo dado se ve que la representación geométrica de la correspondencia inversa se obtiene cambiando la dirección de las flechas en la representación geométrica de la correspondencia directa. De ello se desprende que la correspondencia inversa de la correspondencia inversa será la correspondencia directa

$$(q^{-1})^{-1}=q. \tag{1-51}$$

c) Composición de correspondencias

Se llama composición de correspondencias al empleo consecutivo de dos correspondencias.

La composición de correspondencias es una operación con tres conjuntos X, Y, Z, en los cuales están definidas dos correspondencias

$$\left.\begin{array}{ll} q=(X, Y, Q), & Q \subseteq X \times Y; \\ p=(Y, Z, P), & P \subseteq Y \times Z, \end{array}\right\} \tag{1-52}$$

por cierto, el dominio de valores de la primera correspondencia coincide con el dominio de definición de la segunda

$$\mathrm{Pr}_2 Q=\mathrm{Pr}_1 P. \tag{1-53}$$

La primera correspondencia define para cualquier $x \in \mathrm{Pr}_1 Q$ cierto elemento $y \in Y$ que además puede ser no único. De acuerdo con la definición de la operación de composición de correspondencias, ahora hace falta definir la $z \in Z$ para la $y \in Y$ hallada, empleando la segunda correspondencia. De este modo, la composición

de correspondencias compara con cada elemento x del dominio de definición de la primera correspondencia $\mathrm{Pr}_1 Q$, uno o varios elementos z del dominio de valores de la segunda correspondencia $\mathrm{Pr}_2 P$.

Vamos a designar por $q(p)$ la composición de correspondencias q y p, y por $Q \circ P$, la gráfica de la composición de correspondencias. Entonces la composición de correspondencias (1-52) se escribe en la forma

$$q(p) = (Y, Z, Q \circ P), \quad Q \circ P \subseteq X \times Z. \tag{1-54}$$

Ejemplo 1-16. Si q es la correspondencia que determina la distribución de los chóferes por vehículos automóviles y p, la correspondencia que determina la distribución de vehículos por rutas, entonces la correspondencia $q(p)$ es la correspondencia que define la distribución de chóferes por rutas.

Como es lógico, la operación de composición puede extenderse asimismo a un número de correspondencias mayor de dos.

1-5. REFLEJOS Y FUNCIONES

a) Reflejos y sus propiedades

Sean X e Y ciertos conjuntos siendo $\Gamma \subseteq X \times Y$ y $\mathrm{Pr}_1 \Gamma = X$. La tríada de conjuntos (X, Y, Γ) define cierta correspondencia que posee la propiedad de que su dominio de definición $\mathrm{Pr}_1 \Gamma$ coincide

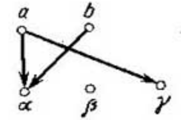


Fig. 1-11. Representación geométrica del reflejo

con el dominio de partida, es decir, con X, y, por lo tanto, dicha correspondencia está definida en todas partes en X. En otras palabras, para cada $x \in X$ existe una $y \in Y$ tal que $(x, y) \in \Gamma$. Tal correspondencia definida en todas partes se denomina *reflejo* de X en Y y se expresa por

$$\Gamma : X \to Y. \tag{1-55}$$

A menudo por la palabra "reflejo" se comprende el reflejo unívoco. Sin embargo, nosotros no vamos a atenernos a esta regla y consideraremos que el reflejo Γ pone en correspondencia con cada elemento $x \in X$ cierto subconjunto

$$\Gamma x \subseteq Y, \tag{1-56}$$

llamado imagen del elemento x. La ley conforme a la cual se realiza la correspondencia se define por el conjunto Γ.

Ejemplo 1-17. Si en el ejemplo 1-14 se excluye del análisis el chófer c, se obtiene el reflejo $\Gamma: X \to Y$, en el cual $X = \{a, b\}$ es el conjunto de los chóferes; $Y = \{\alpha, \beta, \gamma\}$, el conjunto de los vehículos automóviles, y $\Gamma = \{(a, \alpha), (a, \gamma), (b, \alpha)\}$, la distribución de los chóferes por vehículos. En la figura 1-11 se muestra la representación geométrica de este reflejo.

Examinemos algunas propiedades del reflejo. Sea $A \subseteq X$. Para cualquier $x \in A$ el conjunto $\Gamma x \subseteq Y$ será la imagen de x. El total de los elementos de Y que constituyen imágenes de Γx para todas las $x \in A$ se denomina imagen del conjunto A y se designa por ΓA. Según esta definición

$$\Gamma A = \bigcup_{x \in A} \Gamma x. \tag{1-57}$$

Si A_1 y A_2 son subconjuntos de X se cumple que

$$\Gamma (A_1 \cup A_2) = \Gamma A_1 \cup \Gamma A_2. \tag{1-58}$$

En efecto,

$$\Gamma (A_1 \cup A_2) = \bigcup_{x \in A_1 \cup A_2} \Gamma x = \left(\bigcup_{x \in A_1} \Gamma x \right) \cup \left(\bigcup_{x \in A_2} \Gamma x \right) = \Gamma A_1 \cup \Gamma A_2.$$

Sin embargo, la relación

$$\Gamma (A_1 \cap A_2) = \Gamma A_1 \cap \Gamma A_2 \tag{1-59}$$

sólo es válida en caso de que el reflejo sea unívoco. Para demostrar esto realizaremos el fraccionamiento de los conjuntos A_1 y A_2 (fig. 1-12) del tipo

$$A_1 = X_1 \cup X_0, \quad A_2 = X_2 \cup X_0,$$

donde $X_0 = A_1 \cap A_2$.

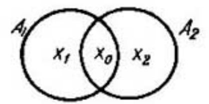


Fig. 1-12. Fraccionamiento de conjuntos intersecados

Por cierto, los conjuntos X_1, X_2, X_0 serán fraccionamientos del conjunto $A_1 \cup A_2$. Aunque dichos conjuntos no son incidentes, sus imágenes, sin embargo, pueden tener elementos comunes en caso de no ser unívoco el reflejo. Por consiguiente,

$$\Gamma A_1 \cap \Gamma A_2 = \Gamma (X_1 \cup X_0) \cap \Gamma (X_2 \cup X_0) =$$
$$= (\Gamma X_1 \cap \Gamma X_2) \cup (\Gamma X_1 \cap \Gamma X_0) \cup (\Gamma X_2 \cap \Gamma X_0) \cup \Gamma X_0.$$

En esta relación se ve que (1-59) sólo se cumple en el caso de que

$$\Gamma X_i \cap \Gamma X_k = \varnothing; \quad i, k \in \{0, 1, 2\}; \quad i \neq k, \tag{1-60}$$

es decir, cuando el reflejo es unívoco. Ahora bien, en el caso general

$$\Gamma X_0 = \Gamma (A_1 \cap A_2) \subseteq \Gamma A_1 \cap \Gamma A_2. \qquad (1\text{-}61)$$

Las relaciones obtenidas se generalizan también con facilidad para un número mayor de conjuntos A_i. Así pues, si $A_1, \ldots, A_n$ son subconjuntos de X, entonces

$$\Gamma \left(\bigcup_{i=1}^{n} A_i \right) = \bigcup_{i=1}^{n} \Gamma A_i; \qquad (1\text{-}62)$$

$$\Gamma \left(\bigcap_{i=1}^{n} A_i \right) \subseteq \bigcap_{i=1}^{n} \Gamma A_i. \qquad (1\text{-}63)$$

Por cuanto el reflejo constituye un caso particular de correspondencia, son válidos para el reflejo los conceptos de reflejo inverso y composición de reflejos, análogos a los conceptos introducidos al estudiar las correspondencias.

b) Reflejos presentados en un conjunto

El caso en que coinciden los conjuntos X e Y es un importante caso particular de reflejo. En tal caso el reflejo $\Gamma: X \to X$ va a representar el reflejo del conjunto X sobre sí mismo y va a definirse por el par

$$(X, \Gamma), \qquad (1\text{-}64)$$

en el cual $\Gamma \subseteq X^2$. La teoría de los grafos, cuyos elementos serán examinados en el capítulo 2, se ocupa del estudio detallado de tales reflejos. Aquí sólo nos referiremos a ciertas operaciones con semejantes reflejos.

Sean Γ y Δ los reflejos del conjunto X en X. Vamos a denominar composición de dichos reflejos el reflejo $\Gamma\Delta$ que en concordancia con la regla expuesta en el § 1-4 se define de la manera siguiente:

$$(\Gamma\Delta)\, x = \Gamma (\Delta x). \qquad (1\text{-}65)$$

En el caso particular en que $\Delta = \Gamma$, obtenemos los reflejos

$$\Gamma^2 x = \Gamma (\Gamma x) \qquad (1\text{-}66)$$

$$\Gamma^3 x = \Gamma (\Gamma^2 x), \text{ etc.} \qquad (1\text{-}67)$$

De este modo, en el caso general para cualquier $s \geqslant 2$ se verifica

$$\Gamma^s x = \Gamma \left(\Gamma^{s-1} x \right). \qquad (1\text{-}68)$$

Mediante una definición especial introduzcamos la relación

$$\Gamma^0 x = x. \qquad (1\text{-}69)$$

Ello da la posibilidad de extender la relación (1-68) también a las s negativas. Efectivamente, según (1-68)

$$\Gamma^0 x = \Gamma(\Gamma^{-1}x) = \Gamma\Gamma^{-1}x = x. \tag{1-70}$$

Esto significa que $\Gamma^{-1}x$ constituye un reflejo inverso. Entonces

$$\Gamma^{-2}x = \Gamma^{-1}(\Gamma^{-1}x), \tag{1-71}$$

etc.

Ejemplo 1-18. Sea X un conjunto de personas. Para cada persona $x \in X$, designemos por Γx el conjunto de sus hijos. Entonces $\Gamma^2 x$ es el conjunto de los nietos de x; $\Gamma^3 x$, el conjunto de los biznietos de x; $\Gamma^{-1}x$, el conjunto de los padres de x, etc.

Representando las personas por puntos y trazando las flechas que van de x a Γx, obtenemos el árbol genealógico que se muestra a título de ejemplo en la figura 1-13.

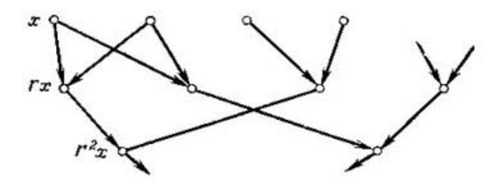


Fig. 1-13. Arbol genealógico

Ejemplo 1-19. Analicemos un juego de ajedrez. Designemos por x cierta posición (ubicación de las figuras en el tablero), que puede crearse durante el juego, y por X, el conjunto de las posiciones posibles. Entonces, Γx va a significar para cualquier $x \in X$ el conjunto de posiciones que pueden obtenerse a partir de x haciendo una jugada y cumpliendo las reglas del juego. En tal caso,

$\Gamma x = \varnothing$, si x es una posición de jaque mate o de tablas;

$\Gamma^3 x$ es el conjunto de las posiciones que pueden obtenerse a partir de x con tres jugadas;

$\Gamma^{-1}x$ es el conjunto de las posiciones a partir de las cuales puede obtenerse la posición dada con una jugada.

Para los reflejos presentados en un conjunto, con frecuencia se utilizan algunas otras denominaciones que se encontrarán más adelante en nuestra obra.

Así pues, si los elementos $x \in X$ representan los estados de un sistema dinámico, el reflejo Γx puede considerarse como el conjunto de los estados a los cuales puede pasar el sistema partiendo de un estado dado. En este caso lo más apropiado es utilizar el término transformación del estado del sistema dinámico. Para designar ciertos tipos especiales de reflejos presentados en un mismo conjunto, también se emplea el término *relación*.

c) Función, funcional y operador

Examinemos cierto reflejo

$$f : X \rightarrow Y. \qquad (1\text{-}72)$$

Este reflejo se denomina *función* si es unívoco, es decir, si para cualesquiera pares $(x_1, y_1) \in f$ y $(x_2, y_2) \in f$ de $x_2 = x_1$ se desprende que $y_2 = y_1$.

De la definición de reflejo y de los ejemplos dados anteriormente se deduce que los elementos de los conjuntos X e Y pueden ser objetos de cualquier naturaleza. No obstante, en los problemas de cibernética presentan gran interés los reflejos que son unívocos y cuyo conjunto de valores es el conjunto de los números reales R. El reflejo unívoco f definido por (1-72) se denomina función con valores reales si $Y \subseteq R$.

Ejemplo 1-20. De una ciudad a otra se puede viajar por tren, autobús o avión. El precio del pasaje es, respectivamente, de 7, 9 y 12 rublos. El precio del pasaje en este ejemplo puede representarse como función del medio de transporte. Examinemos para ello los conjuntos

$$X = \{\text{tr., aut., av.}\}, \quad Y = \{7, 9, 12\}.$$

La función $f: X \rightarrow Y$, obtenida de las condiciones del ejemplo, puede escribirse en forma del conjunto $f\{(\text{tr.}, 7), \text{aut.}, 9), (\text{av.}, 12)\}$.

El valor y en cualquiera de los pares $(x, y) \in f$ se denomina función de la x dada, escribiéndose en la forma $y = f(x)$. Tal notación permite introducir la siguiente definición formal de función:

$$f = \{(x, y) \in X \times Y : y = f(x)\}. \qquad (1\text{-}73)$$

De esta suerte, el símbolo f se utiliza para la definición de la función en dos sentidos:

1) f es el conjunto cuyos elementos son los pares (x, y) que participan en la correspondencia;

2) $f(x)$ es la designación para la $y \in Y$ correspondiente a la $x \in X$ dada.

La definición formal de función a modo de la relación (1-73), permite establecer los métodos de presentación de la función.

1. La enumeración de todos los pares (x, y) que constituyen el conjunto f, como se hizo en el ejemplo 1-20. Tal método de presentación de la función es aplicable si X es un cojunto finito. Para mayor claridad conviene disponer los pares (x, y) en forma tabular.

2. En muchos casos, tanto X como Y constituyen conjuntos de números reales o complejos. En tales casos, muy frecuentemente se comprende por $f(x)$ la fórmula, es decir, la expresión que contiene la relación de las operaciones matemáticas (adición, substracción, división, logaritmación, etc.), que hay que realizar con $x \in X$ para obtener y.

Ejemplo 1-21. Sean $X = Y = R$ y $f = \{(x, y) \in R^2 : y = x^2\}$. Entonces $f(x) = x^2$.

A veces resulta necesario emplear distintas fórmulas para los diferentes subconjuntos del conjunto X de la función. Sean $A_1, \ldots, A_n$ los subconjuntos de X que no se intersecan por pares. Designemos por $f_i(x)$ $(i = 1, \ldots, n)$ la fórmula que define y siendo $x \in A_i$. Entonces la función $f(x)$ va a definirse por la expresión

$$f(x) = \begin{cases} f_1(x) & \text{para } x \in A_1; \\ \cdot \quad \cdot \quad \cdot & \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ f_n(x) & \text{para } x \in A_n. \end{cases} \qquad (1\text{-}74)$$

Así, la función $y = f(x) = |x|$ puede presentarse en la forma

$$y = \begin{cases} x & \text{para } x \geqslant 0; \\ -x & \text{para } x < 0. \end{cases}$$

3. Si X e Y son conjuntos de los números reales, los elementos $(x, y) \in f$ pueden representarse en forma de puntos en el plano R^2. La totalidad de dichos puntos va a constituir la *gráfica* de la función $f(x)$. En la figura 1-14, *a* está representada la gráfica de la función dada en el ejemplo 1-21.

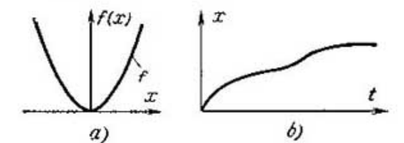


Fig. 1.14. Gráficas de funciones

En los problemas de cibernética con mucha frecuencia se hace necesario operar con funciones del tiempo. Estas funciones definen el reflejo de un número finito o infinito de puntos de cierto intervalo de tiempo T en el conjunto de números reales $X \subseteq R$, lo que puede escribirse en la forma

$$f : T \to X. \qquad (1\text{-}75)$$

Designando por t los elementos del conjunto T, y por x los elementos del conjunto X, obtenemos la función $x = f(t)$, que determina el carácter de la variación del valor de x en el tiempo, como se muestra por ejemplo, en la figura 1-14, *b*. Para simplificar la notación vamos a designar simplemente por $x(t)$ la dependencia de x del tiempo.

Si en la expresión (1-72) $X = U \times V$, entonces llegamos a la función de dos variables u y v, designada por $f(u, v)$, donde $u \in U$ y $v \in V$. La definición formal de la función de dos variables reales será la siguiente:

$$f = \{(u, v, y) \in U \times V \times Y : y = f(u, v)\}. \qquad (1\text{-}76)$$

De manera análoga se definen las funciones de tres o un número mayor de variables.

Por cuanto la función es un caso particular de correspondencia, para ella serán válidos los conceptos de función inversa y composición de las funciones, como los expuestos para la correspondencia. Si f y g son dos funciones en el conjunto R^2, siendo

$$f: X \to Y \quad g: Y \to Z, \tag{1-77}$$

entonces las funciones inversas serán:

$$f^{-1}: Y \to X \quad g^{-1}: Z \to Y. \tag{1-78}$$

La composición de las funciones f y g

$$f \circ g: X \to Z \tag{1-79}$$

define para cada $x \in Z$ una $z \in Z$ que se designa como

$$z = (f \circ g)\, x = g\,[f\,(x)]. \tag{1-80}$$

El concepto de funcional es un concepto más general que el de función. La *funcional* establece la dependencia entre un conjunto de números, por un lado, y cierto conjunto de funciones, por otro lado. Puede servir de ejemplo de funcional la integral definida del tipo

$$J\,(f) = \int_a^b f\,(x)\,dx.$$

Como vemos, la funcional $J(f)$ representa un número dependiente de la función $f(x)$, la cual se elige de cierto conjunto presentado de funciones.

El concepto de operador es un concepto aún más amplio. El *operador* establece la correspondencia entre dos conjuntos de funciones de tal modo que a cada uno de éstos corresponde una función determinada del otro conjunto. Así pues, si designamos por p el operador de diferenciación, la relación entre la derivada $f'(x) = df(x)/dx$ y la función $f(x)$ puede escribirse en forma de la relación operacional

$$f'(x) = p\,[f\,(x)].$$

1-6. RELACIONES

a) Relación como representación del enlace mutuo entre los fenómenos

Hasta ahora hemos considerado los conjuntos simplemente como una reunión de elementos. En los ejemplos dados hemos visto que pueden ser elementos de los conjuntos los objetos y fenómenos más diversos de la naturaleza, la técnica y la sociedad, e incluso

de la vida cotidiana. Estos objetos y fenómenos poseen las más diferentes propiedades. Pero, al estudiar los conjuntos como reunión de elementos nos hemos abstraído de todas las propiedades de dichos elementos salvo una: ser elementos del conjunto considerado.

Tal enfoque resulta sumamente provechoso ya que permite introducir un método único de descripción de los fenómenos de más diverso tipo. Pero este enfoque en muchos casos es a la vez unilateral e insuficiente.

En la naturaleza y la sociedad los diferentes sucesos y fenómenos no existen por sí mismos, sino que están vinculados recíprocamente e influyen unos sobre otros. En el concepto de conjunto no se refleja este enlace mutuo de los fenómenos. Es por ello que el concepto de conjunto requiere su desarrollo ulterior para aproximarse a la descripción de las situaciones reales.

El primer paso en este sentido es la introducción del concepto de relación en el conjunto que constituye el método matemático de expresión del enlace mutuo entre los fenómenos. El concepto de conjunto obtiene su desarrollo posterior en los espacios multidimensionales al estudio de los cuales está dedicado el tercer capítulo.

b) Propiedades de las relaciones

Como ya se ha indicado, el término "relación" se emplea para designar algunos tipos de reflejos dados en un mismo conjunto. Debido al uso de este término es conveniente introducir una serie de símbolos especiales.

Supongamos que el reflejo (X, Γ) es una relación. Examinemos el elemento $y \in \Gamma x$. Diremos que el elemento y se encuentra en relación Γ con el elemento x y lo escribiremos en la forma

$$y \Gamma x. \tag{1-81}$$

Así pues, el símbolo Γ en el ejemplo 1-18 denota la relación "ser hijos de la persona dada".

Nota. Utilizando para el reflejo dado en un conjunto la correlación (1-64), obtenemos que la relación es el par de conjuntos (X, Γ) en el cual $\Gamma \subseteq X^2$. Por cuanto los elementos del conjunto X^2 son pares ordenados puede decirse que la relación es un conjunto de pares ordenados. Como que cada par relaciona entre sí sólo dos elementos del conjunto X^2, esta relación se denomina a veces relación *binaria* o *de dos lugares*.

Puede introducirse un concepto de relación más general denominando relación el par de conjuntos (X, Γ), donde $\Gamma \subseteq X^n$. Los elementos del conjunto X^n son las n-es ordenadas y esto permite denominar *n-aria* o *de n lugares* la relación dada. En particular, el conjunto de las triadas ordenadas puede llamarse relación *ter-*

naria o *de tres lugares*. En lo adelante, sin especificarlo especialmente, con el término "relación" nos referiremos a la relación *binaria*.

Las relaciones se dividen en diferentes clases dependiendo de si poseen o no ciertas propiedades.

Examinemos las seis propiedades principales de las relaciones. Al exponer dichas propiedades vamos a considerar que x, y, z son elementos cualesquiera del conjunto X.

Reflexividad: $x\Gamma x$ es cierto; antirreflexividad: $x\Gamma x$ es falso; simetría: $x\Gamma y \rightarrow y\Gamma x$; antisimetría: $x\Gamma y$ y $y\Gamma x \rightarrow x = y$; asimetría: si $x\Gamma y$ es cierto, entonces $y\Gamma x$ es falso; transitividad: $x\Gamma y$ y $y\Gamma z \rightarrow$ $\rightarrow x\Gamma z$.

Examinemos algunas clases importantes de relaciones, utilizando las propiedades expuestas.

c) Relación de equivalencia

Algunos elementos de un conjunto pueden considerarse equivalentes, cuando, en algunos casos, al someterlos a cierto análisis, uno de ellos puede ser sustituido por otro. Entonces se dice que los elementos dados están en relación de equivalencia.

Ejemplos de relaciones de equivalencia:

la relación "estar en un curso" en el conjunto de los estudiantes de una facultad;

la relación "tener igual resta al dividirse por 3" en el conjunto de los números enteros positivos;

la relación de paralelismo en el conjunto de las rectas de un plano;

la relación de semejanza en el conjunto de los triángulos, etc.

Para formular con exactitud la relación de equivalencia vamos a considerar que el término "relación de equivalencia" se emplea solamente en caso en que se cumplen las tres condiciones siguientes:

1) cada elemento es equivalente a sí mismo;

2) la proposición de que dos elementos son equivalentes no requiere que se precise cuál de los elementos se considera primero y cuál segundo;

3) dos elementos equivalentes a un tercero son equivalentes entre sí.

Adoptemos el símbolo $\equiv$ para denotar equivalencia. Entonces la definición general de equivalencia se obtiene escribiendo las tres condiciones citadas anteriormente en forma de las relaciones siguientes:

1) $x \equiv x$ (reflexividad);

2) $x \equiv y \rightarrow y \equiv x$ (simetría);

3) $x \equiv y$ e $y \equiv z \rightarrow x \equiv z$ (transitividad).

Así pues, la relación Γ se denomina relación de equivalencia si es reflexiva, simétrica y transitiva.

La relación de equivalencia está vinculada estrechamente al fraccionamiento de conjuntos estudiado en el § 1-2. Sea X un conjunto en el que está definida la relación de equivalencia. Por ejemplo, X es el conjunto de estudiantes en un curso y la relación de equivalencia es "estar en un grupo". Llamaremos clase de equivalencia al subconjunto de elementos equivalentes a cierto elemento $x \in X$. Así pues, el grupo en que cursa estudios el estudiante Ivanov será la clase de equivalencia que equivale al estudiante mencionado.

Sea J cierto conjunto de índices. Designemos por $\{A_j \subseteq X : j \in J\}$ el conjunto de clases de equivalencia para el conjunto X. Evidentemente, todos los elementos de una clase de equivalencia son equivalentes entre sí (propiedad transitiva) y cualquier elemento $x \in X$ puede estar en una y sólo una clase. Pero en tal caso X es la reunión de los conjuntos no intersecados A_j y de este modo el sistema completo de clases $\{A_j \subseteq X : j \in J\}$ constituye el fraccionamiento del conjunto X. Así pues, a cada relación de equivalencia en el conjunto X corresponde cierto fraccionamiento de éste en clases A_j.

La relación de equivalencia en el conjunto X y el fraccionamiento del mismo en clases se denominan conjugados si para x e y cualesquiera la relación $x \equiv y$ se cumple cuando, y solo cuando x e y pertenecen a una misma clase A_j de este fraccionamiento. La comparación de los ejemplos de la sección presente con los ejemplos del § 1-2 coadyuvará a esclarecer con más exactitud el vínculo de la relación de equivalencia con el fraccionamiento del conjunto.

En calidad de símbolo general de la relación de equivalencia se utiliza el símbolo $\equiv$ (a veces $\sim$). No obstante, para algunas relaciones de equivalencia particulares se emplean otros símbolos: $=$ para designar igualdad; $\parallel$ para denotar paralelismo; $\rightleftarrows$ para designar equivalencia lógica.

d) Relación de orden

Con frecuencia encontramos relaciones que definen cierto orden de disposición de los elementos de un conjunto. Por ejemplo, distinguimos los conceptos "antes" y "después" en aquellos casos en que los elementos del conjunto son estados de un sistema dinámico. Diferenciamos los conceptos "es menor que" y "es mayor que" y utilizamos los símbolos $>$ o $<$ si los elementos del conjunto son números. Distinguimos los conceptos de conjunto y subconjunto empleando los símbolos $\subseteq$ o $\subset$.

En todos estos casos los elementos del conjunto X o los grupos de elementos pueden disponerse en cierto orden, es decir, puede introducirse la relación de orden en el conjunto X.

Se distingue la relación de orden no estricto para la que se utiliza el símbolo $\leqslant$ (siendo sus casos particulares los símbo-

los $\leqslant$, $\subseteq$) y la relación de orden estricto para la cual se emplea el símbolo $\prec$ (siendo sus casos particulares los símbolos $<$, $\subset$, $\rightarrow$). Describamos estas relaciones enumerando sus propiedades.

Se denomina relación de orden no estricto la que posee las tres propiedades siguientes:

$x \leqslant x$ es cierto (reflexividad);
$x \leqslant y$ e $y \leqslant x \rightarrow x = y$ (antisimetría);
$x \leqslant y$ e $y \leqslant z \rightarrow x \leqslant z$ (transitividad).

Se llama relación de orden estricto a la que tiene las tres propiedades siguientes:

$x \prec x$ es falso (antirreflexividad);
$x \prec y$ e $y \prec x$ se excluyen mutuamente (asimetría);
$x \prec y$ e $y \prec z \rightarrow x \prec z$ (transitividad).

El conjunto X se denomina ordenado si dos elementos cualesquiera x e y del mismo son comparables, es decir, si para ellos

$$x \prec y \text{ o } x = y \text{ o } y \prec x.$$

e) Relación de predominio

En aquellos casos en que X denota un conjunto de personas o de grupos de personas, nos encontramos con una relación que es de predominio. Diremos que x predomina sobre y y escribiremos $x \ll y$ si x supera en algo a y. Así, x puede ser un deportista o equipo deportivo que venció al deportista o equipo y, o una persona que goza de autoridad ante la persona y, o una propiedad que se prefiere a la propiedad y.

Diremos que entre los elementos del conjunto X hay relación de predominio si dichos elementos tienen las dos propiedades siguientes:

1) ningún individuo puede predominar sobre sí mismo, es decir, $x \gg x$ es falso (antirreflexividad);

2) en cada par de individuos es seguro que un individuo predomina sobre otro, o sea, $x \ll y$ e $y \ll x$ se excluyen mutuamente (asimetría).

Respecto al predominio la propiedad de transitividad no se cumple. En efecto, si en unas competencias el equipo x venció al y y el y derrotó al z, de esto no se deduce que el equipo x vencerá necesariamente al z.

1-7. ALGUNOS CONCEPTOS DE ALGEBRA SUPERIOR

a) Grupos, anillos y campos

Las operaciones algebraicas (adición, multiplicación y división) que fueron introducidas inicialmente para los números racionales, durante el proceso de desarrollo de las matemáticas fueron

extendidas a otras nociones diferentes: números complejos, vectores, matrices, etc. Las reglas de ejecución de estas operaciones son distintas para las diferentes nociones. No obstante, dichas operaciones tienen propiedades comunes cuyo conocimiento permite determinar si es posible o no emplearlas para cualquier tipo concreto de nociones. El establecimiento de tales propiedades conduce a los conceptos de operación, grupo, anillo y campo algebraicos.

Sea X cierto conjunto. Se dice que en el conjunto X está presentada una *operación algebraica* si con cada par ordenado $(a, b) \in X^2$ está puesto en correspondencia de manera unívoca un elemento determinado c que pertenece al mismo conjunto X. La operación definida de este modo se denomina *multiplicación* o *adición* y se escribe en la forma

$$c = ab \quad \text{o} \quad c = a + b. \tag{1-82}$$

La operación algebraica se denomina asociativa si para $a, b, c \in X$ cualesquiera se cumple la correlación

$$(ab)\,c = a\,(bc) \quad \text{o} \quad (a + b) + c = a + (b + c).$$

Se denomina *semigrupo* el conjunto X con una operación asociativa presentada en el mismo (llamémosla por ahora multiplicación). El semigrupo se denomina *grupo* si:

1) en el conjunto X existe un elemento e tal que para cualquier $a \in X$ se verifica

$$ae = ea = a; \tag{1-83}$$

2) para cualquier $a \in X$ existe un elemento a^{-1} tal que

$$aa^{-1} = a^{-1}a = e. \tag{1-84}$$

El elemento e se denomina *unidad* del *grupo* y el elemento a^{-1}, *inverso* de a. Si la operación definida en el grupo se denomina adición, entonces el elemento e se llama *cero* del grupo y se designa por el símbolo 0 y el a^{-1} se denomina *opuesto* de a y se designa $-a$. Los elementos 0 y $-a$ satisfacen las relaciones

$$a + 0 = 0 + a = a, \quad a + (-a) = 0. \tag{1-85}$$

El grupo se denomina *finito* si X es un conjunto finito. Son ejemplos de grupos: el conjunto de los números enteros respecto a la operación de adición; el conjunto de todos los números racionales diferentes de cero respecto a la operación de multiplicación; el conjunto de todos los vectores en el plano respecto a la operación de adición vectorial. Sin embargo, los números enteros no forman grupo respecto a la operación de multiplicación ya que para un número entero distinto de ± 1 no existe un número entero inverso del mismo. El conjunto $\{1, -1, i, -i\}$, donde $i = \sqrt{-1}$ puede servir de ejemplo de grupo finito respecto a la operación de multiplicación.

Si en el conjunto X están definidas a la vez dos operaciones algebraicas, la adición y la multiplicación, siendo conmutativa la operación de adición ($a + b = b + a$), y estando la operación de multiplicación vinculada a la de adición mediante las leyes distributivas

$$a(b + c) = ab + ac \quad \text{y} \quad (b + c)a = ba + ca, \tag{1-86}$$

entonces tal conjunto se denomina *anillo*. En el anillo pueden faltar la unidad y los elementos inversos. Si en él hay la unidad, entonces se denomina *anillo con unidad*. Son anillos los conjuntos de todos los números enteros, racionales, reales y complejos, respecto a las operaciones comunes de adición y multiplicación.

Se denomina *campo* al anillo en el que para elementos cualesquiera $a \neq 0$ y b existe precisamente un elemento x tal que $ax = b$. El elemento x se denomina cociente de la división del elemento b por el a y se designa por $x = b/a$. Sirven de ejemplos de campos el anillo de todos los números racionales, el anillo de todos los números reales y el anillo de todos los números complejos.

b) Isomorfismo. Homomorfismo. Simulación

La simulación de sistemas y situaciones existentes en el mundo real juega gran papel en las investigaciones científicas y prácticas. La esencia de la simulación consiste en establecer una relación de equivalencia entre dos sistemas, cada uno de los cuales puede existir en realidad o ser abstracto. Si el primero resulta más sencillo para la investigación que el segundo, es posible juzgar sobre las propiedades del segundo sistema observando el comportamiento del primero. En este caso el sistema empleado para la investigación se denomina *modelo*.

El modelo se denomina *isomórfico* (de igual forma) si entre el modelo y el sistema real se observa una correspondencia total de elementos. Dicha correspondencia tiene lugar entre el negativo y la imagen obtenida del mismo, el dibujo y la pieza elaborada según éste, los procesos en un sistema real y la solución de la ecuación que describe su comportamiento.

Sin embargo, en muchos casos los modelos isomórficos resultan demasiado complicados e incómodos para su utilización práctica. Resultan más apropiados aquellos que permiten juzgar solamente acerca de los aspectos esenciales del comportamiento de los sistemas reales sin detallarlos. Puede servir de ejemplo el siguiente modelo: el mapa geográfico con respecto al sector de la superficie terrestre representada en éste.

Se llaman *homomorfos* los modelos, algunos de cuyos elementos sólo corresponden a grandes partes del sistema real y en los cuales falta la correspondencia total entre los elementos del modelo y del sistema.

Puede darse una definición matemática rigurosa del isomorfismo y homomorfismo en términos de la teoría de los grupos.

Supongamos que los conjuntos X e Y son grupos. Si entre los elementos de dichos grupos se ha establecido una correspondencia biunívoca, según la cual para elementos cualesquiera $a, b \in X$ y elementos correspondientes a ellos $a', b' \in Y$, al elemento $c = ab$ va a corresponder el elemento $c' = a'b'$, denominándose a tales grupos isomórficos.

Los grupos isomórficos pueden diferir unos de otros sólo por la naturaleza de sus elementos y quizás por la denominación de las operaciones definidas en el grupo. Pero todas las propiedades de los grupos isomórficos entre sí que se derivan de las propiedades de las operaciones definidas en ellos y que no dependen de la naturaleza de los elementos del grupo son iguales.

En los grupos homomórficos la correspondencia entre grupos es unilateral. Se dice que el grupo X *está reflejado homomórficamente* en el grupo Y si a cada elemento del grupo X corresponde un elemento determinado unívocamente del grupo Y, y si a los elementos $a, b \in X$ corresponden los elementos $a', b' \in Y$, entonces al elemento $ab = c$ corresponde el elemento $c' = a'b'$. En lo general, en caso de homomorfismo del tipo $X \to Y$ pueden pasar al elemento considerado del grupo Y diversos elementos del grupo X y puede asimismo no pasar ninguno.

PROBLEMAS PARA EL CAPITULO 1

1-1. ¿Qué representa el conjunto $Y \setminus X$ del ejemplo 1-1?

1-2. Designar mediante rayado el conjunto $Y \setminus X$ del ejemplo 1-3.

1-3. Sea R un conjunto de números reales y

$$X = \{x \in R : 0 \leqslant x \leqslant 1\}, \quad Y = \{y \in R : 0 \leqslant y \leqslant 2\}.$$

¿Qué representan los conjuntos $X \cup Y$, $X \cap Y$, $X \setminus Y$?

1-4. Dibujar las figuras que representan los conjuntos $A = \{(x, y) \in R^2 : : x^2 + y^2 \leqslant 1\}$ y $B = \{(x, y) \in R^2 : x^2 + (y - 1)^2 \leqslant 1\}$.

¿Qué figuras representan los conjuntos $A \cup B$, $A \cap B$, $R^2 \setminus A$?

1-5. Demostrar las identidades $X \cap \varnothing = \varnothing$; $X \cup \varnothing = X$, $I \cap X = X$ $I \cup X = I$ utilizando la correlación (1-34).

1-6. Representar en el plano real $R^2 = R \times R$ los conjuntos $X \times Y$ e $Y \times X$ del problema 1-3.

1-7. Representar geométricamente los conjuntos $A \times R$ y $R \times A$, donde $A = [2, 3]$.

1-8. Hallar $\mathrm{Pr}_1 M$ y $\mathrm{Pr}_2 M$ para el conjunto $M = \{x, y\} \in R^2 : (x - 2)^2 + + y^2 = 1\}$.

1-9. Sea $I = \{x_1, x_2, x_3\}$ un conjunto universal y $X = \{x_1, x_2\}$; $Y = \{x_2, x_3\}$; $Z = \{x_3\}$ sus subconjuntos. Determinar enumerando los conjuntos siguientes:

$$X \times X;\ Z \times Z;\ X \times Y;\ Y \times X;\ X \times Y \cap Y \times X;\ X \times Y \cup Y \times X.$$

1-10. Sean X, Y, Z los subconjuntos del conjunto R^2 iguales a $X = \{(x, y) : : x \geqslant 0\}$; $Y = \{(x, y) : y \geqslant 0\}$; $Z = \{(x, y) : x + y \geqslant 1\}$. Representar geométrica-

mente los conjuntos

$$X;\ Y;\ Z;\ X\cup Y;\ \overline{X\cup Y};\ X\cap Y;\ \overline{X\cap Y};\ X\cap Z;\ \overline{X\cap Z};$$

$$X\cap Y\cap Z;\ X\cap Y\cap \bar{Z}.$$

1-11. ¿A qué es igual el conjunto $X\times Y$ si X e Y son subconjuntos del conjunto R^2 y $X=\{(x, y): 2x+y=1\}$; $Y=\{(x; y): x-y=0\}$?

1-12. Escribir en el ejemplo 1-13 todas las 16 correspondencias. Determinar Pr_1Q y Pr_2Q para cada una de ellas.

1-13. Hallar la función inversa f^{-1} para la función f del ejemplo 1-21.

1-14. Sean f y g las funciones en el conjunto R^2 iguales a $f=\{(x, y): y: x^2\}$; $g=\{(y, z): z=\text{sen}\, y\}$. Hallar la composición de estas funciones $f\circ g$.

Capítulo segundo

FUNDAMENTOS DE LA TEORIA DE LOS GRAFOS

2-1. DEFINICIONES FUNDAMENTALES DE LA TEORIA DE LOS GRAFOS

a) Definición de grafo según la teoría de conjuntos

Se puede obtener una representación clara del grafo si nos imaginamos cierto conjunto de puntos del plano X denominados vértices y el conjunto de segmentos dirigidos U llamados arcos que unen todos o algunos de los vértices. Matemáticamente, el grafo G puede definirse como el par de conjuntos X y U:

$$G = (X, U). \tag{2-1}$$

En la figura 2-1 está representado el **grafo cuyos** vértices son los puntos a, b, c, d, e, g, h, y sus arcos, los segmentos (a, a), (c, b), (c, d), (d, c); (d, d), (c, e), (e, d), (g, h). Son ejemplos de grafos las relaciones de paternidad y maternidad en un conjunto de personas (véase la fig. 1-13), el mapa de los caminos en la comarca, el diagrama de conexiones de aparatos eléctricos, las relaciones de superioridad de algunos participantes de un torneo sobre otros, etc.

A veces resulta cómodo dar otra definición de grafos. Puede considerarse que el conjunto de los arcos dirigidos U que unen los elementos del conjunto X refleja este conjunto en sí mismo. Por eso, puede considerarse presentado el grafo si están dados el conjunto de sus vértices X y el método de reflexión Γ del conjunto X en X. De esta suerte, el grafo G es el par (X, Γ), que consta del conjunto X y el reflejo Γ presentado en este conjunto:

$$G = (X, \Gamma). \tag{2-2}$$

Así, para el grafo representado en la figura 2-1 el reflejo Γ se determina del modo siguiente:

$$a = \Gamma a;\ \ \Gamma b = \varnothing;\ \ \Gamma c = \{b, d, e\};\ \ \Gamma d = \{d, c\};\ \ \Gamma e = d;$$
$$\Gamma g = h;\ \ \Gamma h = \varnothing.$$

Es fácil ver que la definición de grafos dada coincide por completo con la definición de relación en el conjunto.

En ocasiones es conveniente representar los grafos en forma de ciertas matrices, en particular como matrices de adyacencia y de incidencia. Daremos previamente dos definiciones.

Dos vértices x e y se denominan *adyacentes* si son diferentes y existe un arco que va de x a y.

El arco u se llama *incidente* con el vértice x si llega a este vértice o sale del mismo.

Designemos por $x_1, \ldots, x_n$ los vértices del grafo y por $u_1, \ldots, u_m$ sus arcos. Introduzcamos los números:

$$r_{ij} = \begin{cases} 1, & \text{si hay un arco que une el vértice} \\ i, & \text{con el vértice } j; \\ 0, & \text{si no hay tal arco.} \end{cases}$$

La matriz cuadrática $\mathbf{R} = \| r_{ij} \|$ del orden $n \times n$ se denomina *matriz de adyacencia* del grafo,

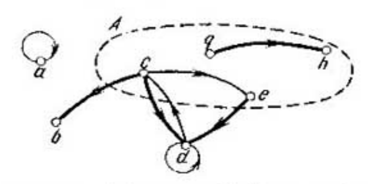


Fig. 2-1. Vista general de un grafo

Introduzcamos más adelante los números

$$s_{ij} = \begin{cases} +1, & \text{si } u_j \text{ sale de } x_i \\ -1, & \text{si } u_j \text{ llega a } x_i; \\ 0, & \text{si } u_j \text{ no es incidente con } x_i. \end{cases}$$

La matriz $\mathbf{S} = \|s_{ij}\|$ de orden $n \times m$ se llama *matriz de incidencia* de los arcos del grafo.

Las matrices de incidencias sólo son aplicables en la forma descrita a los grafos sin lazos. En caso de haber lazos en el grafo hay que desmembrar dicha matriz en dos semimatrices: positiva y negativa.

Introduzcamos algunos conceptos y definiciones que sirven para describir los distintos tipos de grafos.

Subgrafo G_A del grafo $G = (X, \Gamma)$ se denomina el grafo que incluye sólo una parte de los vértices del grafo G que, junto con los arcos que unen dichos vértices, forman el conjunto A, como por ejemplo, la zona circunscrita con línea de trazos en la figura 2-1. El subgrafo G_A se define matemáticamente de la siguiente manera:

$$G_A = (A, \Gamma_A), \tag{2-3}$$

donde

$$A \subseteq X, \quad \Gamma_A x = (\Gamma x) \cap A. \tag{2-4}$$

El *grafo parcial* G_Δ con respecto al grafo $G=(X, \Gamma)$ se llama al grafo que sólo contiene una parte de los arcos del grafo G, es decir, definido por la condición

$$G_\Delta=(X, \Delta), \tag{2-5}$$

donde

$$\Delta x \subseteq \Gamma x. \tag{2-6}$$

Así pues, en la figura 2-1 el grafo formado por los arcos gruesos es un grafo parcial.

Ejemplo 2-1. Sea $G=(X,\Gamma)$ el mapa de carreteras de la Unión Soviética. Entonces el mapa de carreteras de la región de Tambov constituye un subgrafo y el mapa de carreteras principales del la Unión Soviética, un grafo parcial.

Son también importantes los conceptos de camino y contorno. Anteriormente se dio la definición de arco como el segmento dirigido que une dos vértices. El arco que une los vértices a y b y está dirigido de a a b se designa por $u=(a, b)$.

En el grafo G se denomina *camino* la secuencia de arcos $\mu = (u_1, \ldots, u_k)$ en la cual el final de cada arco anterior coincide con el comienzo del siguiente. El camino μ cuyos vértices consecutivos son $a, b, \ldots, m$ se designa por $\mu=(a, b, \ldots, m)$. Se llama *largo del camino* $\mu=(u_1, \ldots, u_k)$ al número $l(\mu)=k$ igual al número de arcos que constituyen el camino μ. Este puede ser finito o infinito. En caso de ser infinito suponemos que $l(\mu)=\infty$. El camino en el cual ningún arco se encuentra dos veces se denomina *sencillo*. Se llama *elemental* al camino en el cual ningún vértice se encuentra dos veces.

El *contorno* es el camino finito $\mu=(x_1, \ldots, x_k)$ en el que el vértice inicial x_1 coincide con el final x_k. Por cierto, el contorno se denomina *elemental* si todos sus vértices son diferentes (a excepción del inicial y final que coinciden). Se llama *lazo* al contorno de largo unitario formado por el arco de tipo (a, a). Así pues, en la figura 2-1 (e, d, c, b) es un camino, (c, e, d, c), un contorno y (d, d), un lazo.

A veces se analiza el grafo sin tomar en consideración la orientación de sus arcos. En tal caso se denomina grafo *no orientado*. Para éste los conceptos de arco, camino y contorno se sustituyen por los de arista, circuito y ciclo. La *arista* es el segmento que une dos vértices. El grafo en la figura 2-1 tiene ocho arcos y siete aristas. Se denomina *circuito* a una secuencia de aristas. Se llama *ciclo* al circuito finito en el que coinciden los vértices inicial y final.

Está vinculada al concepto de grafo no orientado una importante característica denominada *conexión del grafo*. Se dice que el grafo es conexo si dos de sus vértices cualesquiera pueden conectarse en circuito. Si el grafo G no es conexo puede desmembrarse en subgrafos G_i tales que sean conexos todos los vértices en cada subgrafo y no lo sean los de los diferentes subgrafos.

Tales subgrafos G_i se denominan componentes de conexión del grafo G.

Para determinar la conexión de un grafo orientado no es necesario atender a la orientación de los arcos. El grafo representado en la figura 2-1 no es conexo, pero su subgrafo que consta de los vértices b, c, d, e es conexo. Para el grafo orientado existe el concepto de conexión fuerte. Un grafo es fuertemente conexo si para dos vértices cualesquiera x e y $(x \neq y)$ existe un camino que va de x a y.

El árbol es un importante caso particular de grafo no orientado. Se denomina árbol el grafo conexo finito no orientado desprovisto de ciclos. En la figura 2-2 se muestran ejemplos de árboles.

Si está dado el conjunto de vértices $a, b, c, \ldots$, el árbol puede construirse del modo siguiente. Tomamos como vértice inicial uno de éstos, por ejemplo a, y lo denominamos *raíz* del árbol. Desde

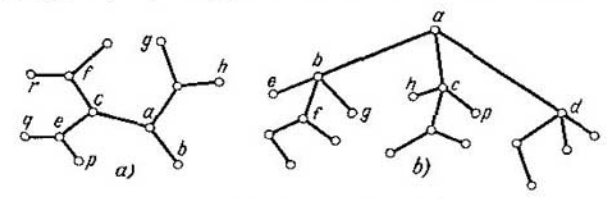


Fig. 2-2. Ejemplos de árboles

este vértice trazamos aristas a los vértices contiguos $b, c, d, \ldots$, desde éstos trazamos las aristas a los vértices vecinos de los mismos $e, f, g, h, \ldots$, etc. De esta manera, el árbol puede construirse añadiendo consecutivamente aristas en sus vértices. Esto ofrece la posibilidad de establecer el vínculo entre el número de vértices y el de aristas del árbol.

El árbol más sencillo consta de dos vértices unidos por una arista. Cada vez que añadimos una arista más en su extremo se adiciona también un vértice. Por lo tanto, un árbol con n vértices tiene $n - 1$ aristas.

b) Relación de orden y relación de equivalencia en un grafo

Como hemos visto, el grafo proporciona una cómoda representación geométrica de las relaciones en un conjunto. Por eso la teoría de los grafos y la teoría de las relaciones en el conjunto se complementan mutuamente.

Vamos a considerar que en el grafo $G = (X, \Gamma)$ se ha introducido una relación de orden si para dos vértices cualesquiera x e y

que satisfagan la condición $x \leqslant y$ existe un camino de x a y. En este caso se dice que el vértice x precede al y y que éste sigue al x.

Mostremos que la definición dada refleja en el grafo todas las propiedades de la relación de orden.

Reflexividad. La condición

$$x \leqslant x \text{ es cierto} \tag{2-7}$$

y denota la equivalencia del vértice a sí mismo, es decir, la condición $x \equiv x$. No obstante, si se quiere, dicha condición puede considerarse como la existencia de un camino de x a x, o sea, como un lazo en el vértice x (fig. 2-3, a).

Transitividad. La condición

$$x \leqslant y, \quad y \leqslant z \rightarrow x \leqslant z \tag{2-8}$$

denota que los vértices x, y, z se encuentran consecutivamente en un mismo camino (fig. 2-3, b)

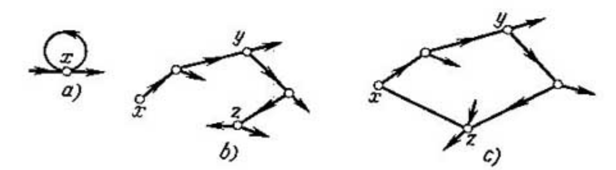


Fig. 2-3. Ilustración de las propiedades de una relación de orden

Asimetría. Demostremos que es cierta la condición

$$x \leqslant y, \quad y \leqslant x \rightarrow x \equiv y. \tag{2-9}$$

El primer miembro de esta expresión indica que existe un camino de x a y, así como también, un camino de y a x. Pero esto significa que en el grafo hay un contorno en el cual están los vértices x e y (fig. 2-3, c).

Del segundo miembro de la condición (2-9) se desprende que los vértices que están en un mismo contorno son equivalentes. Vamos a considerar esta deducción como la definición de equivalencia en el grafo y demostremos que tal definición satisface las tres condiciones de la relación de equivalencia. Las condiciones de reflexividad $x \equiv x$ y simetría $x \equiv y \rightarrow y \equiv x$ son evidentes y se derivan de la definición de equivalencia dada anteriormente. También es evidente la condición de transitividad $x \equiv y$, $y \equiv z \rightarrow$ $\rightarrow x \equiv z$ ya que indica que si en el grafo hay un contorno con los vértices x e y, así como con los vértices y y z, entonces existe asimismo un contorno en el que se encuentran los vértices x y z, (véase la fig. 2-3, c).

De esta suerte, la relación de orden al igual que la relación de equivalencia definen cierto grafo.

En el grafo también puede introducirse la relación de orden estricto. En este caso, para dos vértices cualesquiera x e y que satisfagan la condición $x \prec y$, hay un camino que va de x a y. La condición de transitividad $x \prec y$, $y \prec z \rightarrow x \prec z$ significa, como en el caso anterior, que los vértices x, y, z se encuentran consecutivamente en un mismo camino. La condición de antirreflexividad ($x \prec x$ es falso) indica la ausencia de lazos en el grafo y la condición de asimetría ($x \prec y$, $y \prec x$ se excluyen mutuamente), la falta de contornos.

De este modo, la relación de orden estricto define un grafo sin contornos.

c) Características de los grafos

La resolución de muchos problemas técnicos por los métodos de la teoría de los grafos se reduce a la determinación de unas u otras características de éstos. Aunque en la presente obra es imposible examinar las aplicaciones técnicas de la teoría de los grafos, el conocimiento de las características más relevantes de los grafos, puede resultar de provecho al estudiar otras disciplinas.

Número ciclomático. Sea G un grafo no orientado que tiene n vértices, m aristas y r componentes de conexión. Se denomina número ciclomático de un grafo G el número

$$\nu(G) = m - n + r.$$

Este número tiene un interesante sentido físico: es igual al número máximo de ciclos independientes en el grafo. El número ciclomático puede utilizarse al calcular los circuitos eléctricos para determinar el número de contornos independientes.

Número cromático. Sea p un número natural. El grafo G se denomina p-cromático si pueden colocarse sus vértices con p colores distintos de tal modo que no se coloreen igualmente dos vértices adyacentes cualesquiera. El número menor p con el cual el grafo es p-cromático se denomina número cromático del grafo y se designa por $\gamma(G)$.

Si $\gamma(G) = 2$, el grafo se denomina dicromático. La condición necesaria y suficiente para que el grafo sea dicromático es que no tenga ciclos de largo impar. El número cromático juega un papel importante al resolver el problema de utilizar con máxima economía las células de la memoria durante la programación. Pero salvo en el caso del grafo dicromático su determinación constituye una tarea bastante difícil que con frecuencia requiere el empleo de calculadoras electrónicas.

Conjunto interiormente estable. El conjunto $S \subseteq X$ del grafo $G = (X, \Gamma)$ se denomina interiormente estable si no son adyacentes dos vértices cualesquiera de S, es decir, si para cualquier $x \in S$ se cumple que $\Gamma x \cap S = \varnothing$.

El conjunto interiormente estable que contiene el número mayor de elementos se denomina conjunto interiormente estable máximo, y el número de elementos de este conjunto se llama número de estabilidad interna del grafo G. El conjunto interiormente estable máximo juega gran papelen la teoría de la comunicación.

Conjunto exteriormente estable. El conjunto $T \subset X$ del grafo $G = (X, \Gamma)$ se denomina exteriormente estable si cualquier vértice no pertaneciente a T está conectado por arcos con los vértices que parten de T, es decir, si para cualquier $x \notin T$ se cumple $\Gamma x \cap T \neq \neq \varnothing$.

El conjunto exteriormente estable que contiene el número menor de elementos se llama conjunto exteriormente estable mínimo, y el número de elementos de este conjunto se denomina número de estabilidad externa del grafo G.

2-2. PROBLEMA DEL CAMINO MINIMO

a) Enunciado del problema

Para aplicaciones prácticas tiene gran importancia el problema de encontrar el camino mínimo entre dos vértices de un grafo conexo no orientado. Se reducen a éste muchos problemas de elección de la ruta más económica (desde el punto de vista de la distancia, el tiempo o el costo) en el mapa de carreteras existente y muchos problemas de elección del modo más económico para convertir un sistema dinámico de un estado en otro, etc. En matemáticas se han elaborado diversos métodos para resolver semejantes problemas. Sin embargo, con gran frecuencia los métodos basados en el uso de los grafos son los menos engorrosos.

El problema del camino mínimo en un grafo puede enunciarse en general del modo siguiente. Está dado el grafo no orientado $G = (X, U)$. Se ha asignado a cada arista de este grafo cierto número $l(u) \geqslant 0$ denominado largo de la arista. En casos particulares $l(u)$ puede ser la distancia entre los vértices unidos por la arista u, el tiempo o costo del viaje por esta arista, etc. Entonces, cualquier circuito μ va a caracterizarse por el largo

$$l(\mu) = \sum_{u \in \mu} l(u). \qquad (2\text{-}10)$$

Para dos vértices cualesquiera a y b del grafo G es necesario hallar precisamente aquel camino μ_{ab} cuyo largo total sea mínimo.

Antes de hallar el método general de resolución de este problema examinemos la regla para solucionar los problemas del tipo particular en que el largo de cadaarista es igual a la unidad.

b) Búsqueda del camino mínimo en un grafo con aristas de largo unitario

A veces resulta necesario operar con grafos cuyas aristas tienen un largo igual que se toma como unidad. Los vértices de tal grafo representan habitualmente los estados de algún sistema en el que desde cierto punto de vista son equivalentes todos los tránsitos realizados con un paso. Pongamos un ejemplo de problema que se reduce al examen de un grafo con aristas de largo unitario. El problema dado puede servir de ilustración de los métodos de construcción de grafos para los diferentes casos concretos.

Ejemplo 2-2. Problema de la torre de Hanói. Un tablero tiene tres estacas. En la primera están ensartados m discos cuyo diámetro disminuye de abajo arriba. Se plantea el problema siguiente: haciendo pasar los discos uno por uno,

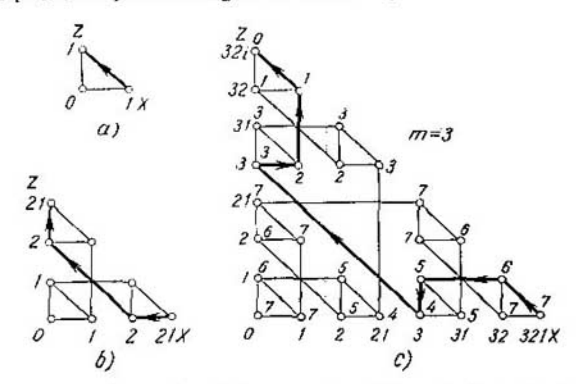


Fig. 2-4. Grafo de las transiciones en el problema de la torre de Hanoi

colocarlos en el mismo orden en la tercera estaca utilizando la segunda en calidad de intermendiaria y cumpliendo la condición de que durante ningún paso el disco mayor quede encima del menor en ninguna de las estacas. Como condición adicional puede exigirse hallar la solución que requiere el menor número de pasos.

Numeremos los discos en orden decreciente de diámetros: m, $m-1$, ..., 1. Designemos por X, Y, Z los conjuntos de discos ensartados, respectivamente, en la primera, segunda y tercera estacas en cualquiera de los pasos. Al hacerlo es suficiente indicar solamente los conjuntos X y Z, ya que el conjunto Y se obtiene como complemento de los conjuntos X y Z hasta el número total de discos. Cada uno de los conjuntos X o Z puede ser una de las combinaciones de discos siguientes: 0, 1, 2, 21, 3, 31, 32, 321, 4, 41, 42, 421, 43, 431, 432, 4321 ... Estas combinaciones pueden representarse con puntos convencionales en los ejes X y Z, como se muestra en la figura 2-4, de suerte que cualquier disposición de los discos se representará con cierto punto en el plano (X, Z). Uniendo dichos puntos por medio de líneas que señalen los desplazamientos posibles de discos en cada paso, obtenemos un grafo no orientado en el que puede hallarse el camino, e incluso el mínimo, para transitar del punto inicial al final del grafo.

Puede realizarse la construcción del grafo pasando de m a $m+1$ discos. Con $m = 1$, los estados posibles van a representarse por el conjunto $\{(1, 0), (0, 0), (0, 1)\}$, al cual corresponde el grafo de la figura 2-4, *a* con el tránsito mínimo del estado inicial $(1, 0)$ al final $(0, 1)$ indicado en ella.

Para obtener una regla general supongamos que ya está construido un grafo para el caso de m discos que llamaremos grafo m. Es fácil cerciorarse de que este grafo tendrá la forma del triángulo representado en la figura 2-5, *a*, aunque sus conexiones internas nos son desconocidas por el momento. ¿Cuál será entonces el grafo para $m+1$ discos?

Notemos que el disco con el número $m+1$ sólo puede ocupar la posición inferior en cualquiera de las estacas. Los m discos restantes pueden desplazarse de cualquier manera conforme al grafo m sin cambiar la posición del disco

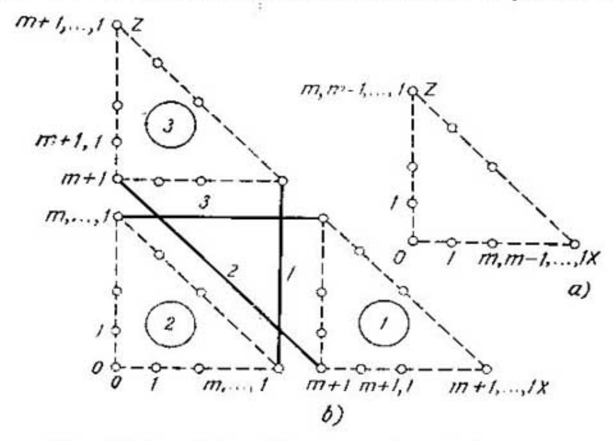


Fig. 2-5. Transición del grafo m al grafo $(m+1)$

$m+1$. Por tanto, el grafo para $m+1$ discos (fig. 2-5, *b*) constará de tres grafos m designados por cifras (encerradas en círculos) *1*, *2* y *3* que significan el número de la estaca en la que se halla el disco $m+1$. Sólo quedan por aclarar las transiciones de un grafo m a otro que corresponden al desplazamiento del disco $m+1$.

Unicamente puede pasarse el disco $m+1$ a una estaca libre y esto sólo es posible en caso de que en una de ellas estén colocados todos los discos menores m. Según esto, en la figura 2-5, *b*, se muestran las transiciones posibles del disco $m+1$ designadas por las cifras *1*, *2* y *3* que denotan el número de la estaca en la que se hallan los m discos menores.

El diagrama mostrado en la figura 2-5, *b* da el principio general de transición del grafo m al grafo $m+1$. En la figura 2-4, *b* y *c* se brindan los grafos de las transiciones para dos y tres discos.

Pasemos al problema de hallar en el grafo el camino mínimo que une el vértice inicial con el final. Por cuanto los grafos considerados son relativamente sencillos, no es difícil hallar el camino mínimo simplemente seleccionando los caminos posibles. No obstante, debe hallarse un método sistemático para los grafos complicados.

La regla general para hallar el camino mínimo en un grafo consiste en atribuir a cada vértice x_i un índice λ_i, igual al largo del camino mínimo desde el vértice dado hasta el vértice final. La asignación de los índices a los vértices en el caso del grafo con aristas de largo unitario se realiza en el orden siguiente:

1) se atribuye el índice 0 al vértice final x_0;

2) se asigna el índice 1 a todos los vértices desde los cuales va la arista hasta el vértice final;

3) se atribuye el índice $\lambda_i + 1$ a todos los vértices que aún no tienen índices y desde los cuales la arista va hasta el vértice con el índice λ_i. Se continúa este proceso hasta que no se marque el vértice inicial. Al terminar de marcar, el índice del vértice inicial será igual al largo del camino mínimo. Hallaremos el mismo camino mínimo si vamos a movernos desde el vértice inicial en la dirección en que disminuyen los índices.

En la figura 2-4, *c* se muestra un ejemplo de marcado y determinación del camino mínimo para $m = 3$.

Señalemos que el método descrito para determinar el camino mínimo es un caso particular de búsqueda de la solución óptima por el método de programación dinámica. Por eso, después de estudiar la programación dinámica será conveniente volver a examinar dicho ejemplo.

c) Búsqueda del camino mínimo en un grafo con aristas de largo arbitrario

La tarea de atribuir índices numéricos a los vértices del grafo se dificulta si las aristas del grafo tienen un largo arbitrario. La dificultad se debe a que en un grafo complejo el camino que pasa por el número menor de vértices, tiene con frecuencia mayor largo que algunos caminos de rodeo. Así pues, en el grafo de la figura 2-6 que representa un mapa de carreteras el camino directo del vértice marcado con un asterisco al vértice final tiene el largo $l = 12$, mientras que el camino indirecto a través del vértice marcado con un triángulo tiene el largo $l = 10$.

El proceso de asignación de índices a los grafos de este tipo consiste en lo siguiente:

1. Cada vértice x_i se marca con el índice λ_i. Inicialmente se atribuye el índice $\lambda_0 = 0$ al vértice final x_0. Para los demás vértices suponemos previamente que $\lambda_i = \infty$ $(i \neq 0)$.

2. Buscamos un arco (x_i, x_j) tal que para él $\lambda_j - \lambda_i > l(x_i, x_j)$ y sustituimos el índice λ_j por el índice $\lambda_j' = \lambda_i + l(x_i, x_j) < \lambda_j$. Continuamos este proceso de sustitución de los índices mientras quede un arco siquiera para el cual se pueda disminuir λ_j.

Apuntemos una propiedad importante que van a poseer los índices atribuidos a los vértices. Sea x_p un vértice arbitrario. Durante el proceso de asignación de índices considerado, el índice λ_p

disminuye monótonamente. Sea x_q el último vértice que sirve para su disminución. Entonces $\lambda_p = \lambda_q + l\ (x_q,\ x_p)$. Por tanto, para el vértice arbitrario x_p con índice λ_p se encuentra el vértice x_q unido a x_p mediante una arista tal que $\lambda_p - \lambda_q = l(x_q,\ x_p)$.

Esta propiedad permite formular la siguiente regla para hallar el camino mínimo.

Sea $x_n = a$ el vértice inicial con el índice χ_n. Hallemos el vértice x_{p_1} tal que $\lambda_n - \lambda_{p_1} = l\,(x_{p_1},\ x_n)$. Hallemos después el vértice x_{p_2} tal que $\lambda_{p_1} - \lambda_{p_2} = l\,(x_{p_2},\ x_{p_1})$, etc., hasta que no lleguemos al vértice final $x_{p_{k+1}} = x_0 = b$. El camino $\mu_0 = (x_n,\ x_{p_1},\ \ldots,\ x_{p_k},\ x_0)$ de largo λ_n es el camino más corto.

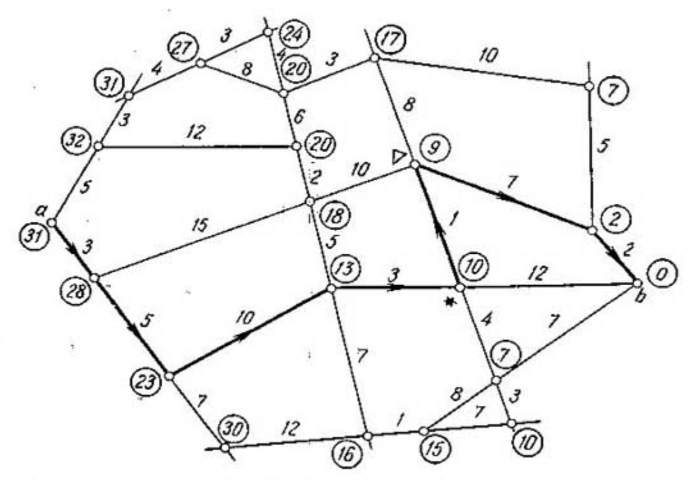


Fig. 2-6. Mapa de carreteras

Para la demostración examinaremos un camino arbitrario de a a b: $\mu = (x_n,\ x_{k_1},\ \ldots,\ x_{k_s},\ x_0)$. Su largo será $l(\mu)$. Conforme a la regla de ordenación de los índices se cumplirán las siguientes desigualdades:

$$\left.\begin{array}{l} \lambda_n - \lambda_{k_1} \leqslant l\,(x_n,\ x_{k_1}); \\ \lambda_{k_1} - \lambda_{k_2} \leqslant l\,(x_{k_1},\ x_{k_2}); \\ \cdots\cdots\cdots \\ \lambda_{k_s} - 0 \leqslant l\,(x_{k_s},\ x_0). \end{array}\right\} \qquad (2\text{-}11)$$

Sumando miembro a miembro estas desigualdades encontramos que para cualquier camino μ se cumple

$$\lambda_n - 0 \leqslant l\,(\mu). \qquad (2\text{-}12)$$

Como que para el camino μ_0 se satisface la condición $\lambda_n = l(\mu_0)$ éste es mínimo.

El método de búsqueda del camino más corto está ilustrado mediante el ejemplo del mapa de carreteras representado en forma de grafo en la figura 2-6. Las cifras en las aristas señalan el tiempo de viaje por cada una de las carreteras. Los índices de los vértices muestran el tiempo de viaje desde el vértice dado hasta el vértice final.

d) Construcción del grafo de largo mínimo

Tiene gran importancia práctica el siguiente problema que puede enunciarse como el problema del trazado de carreteras. Hay varias ciudades $a, b, c \ldots$, que es necesario comunicar entre sí mediante una red de carreteras. Para cada par de ciudades (x, y) se conoce el costo de construcción $l(x, y)$ de la carretera que las une.

El problema consiste en construir la más barata de las redes de carreteras posibles. En lugar de la red de carreteras puede considerarse la de líneas de transporte de energía eléctrica, la de oleoductos, etc. Denominando largo de la arista (x, y) a la magnitud $l(x, y)$ en el grafo que representa la red de carreteras, llegamos al problema de la construcción del grafo de largo mínimo. Por eso, en lo adelanté vamos a considerar el largo de las aristas del grafo en vez del costo de las carreteras.

Si existen sólo tres vértices a, b, c es suficiente construir uno de los circuitos de unión *abc*, *acb* y *bac*, por cierto, si *bc* es la arista más larga entonces hay que excluir precisamente la misma construyendo el circuito *bac*.

El grafo de largo mínimo es siempre un árbol ya que si contuviera un ciclo podría eliminarse una de las aristas de éste y los vértices quedarían unidos todavía. Por consiguiente, para unir n vértices hay que construir $n - 1$ aristas.

Mostremos que puede construirse el grafo de largo mínimo empleando la regla siguiente. Ante todo unimos dos vértices con la arista de unión más corta u_1. En cada uno de los pasos siguientes añadimos la arista más corta de las aristas u_i que no forma ningún ciclo al unirse a las aristas ya existentes. Si hay varias aristas de igual largo elegimos cualquiera de ellas. Vamos a denominar *árbol económico* a cada árbol Q construido de esta manera. Su largo es igual a la suma de las aristas por separado:

$$l(Q) = l(u_1) + \ldots + l(u_{n-1}). \qquad (2\text{-}13)$$

Mostremos que ningún otro árbol que une los mismos vértices puede tener un largo menor que el del árbol económico Q. Sea P el árbol de largo mínimo que une los vértices considerados y Q cualquier árbol económico. Supongamos que las aristas $u_1, u_2, \ldots, u_{n-1}$ están numeradas en el mismo orden en el cual se unieron para construir Q, es decir, satisfacen la condición $l(u_k) \leqslant l(u_{k+1})$.

Si el árbol P no coincide con el Q, este último tiene por lo menos una arista que no pertenece a P. Sea $u_i = (a, b)$ la primera de estas aristas y $L(a, b)$ el circuito del grafo P que une los vértices a y b, como lo muestra por ejemplo la figura 2-7. Si añadimos a P la arista u_i obtenemos un ciclo, y como Q no tiene ciclos, debe formar parte de éste por lo menos una arista no perteneciente a Q. Sea ésta u'_i. Eliminándola, obtenemos el árbol P' con el mismo número de vértices que P, cuyo largo es:

$$l(P') = l(P) + l(u_i) - l(u'_i). \tag{2-14}$$

Como que el grafo P tiene el largo mínimo

$$l(u_i) \geqslant l(u'_i). \tag{2-15}$$

Pero u_i era la arista de largo mínimo con la que no se obtienen ciclos al añadirla a las aristas $u_1, u_2, \ldots, u_{i-1}$. Como que al adi-

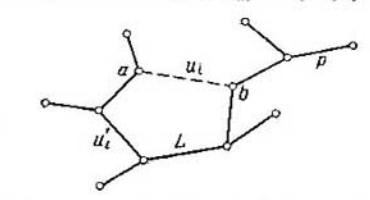


Fig. 2-7. Para la construcción del árbol de largo mínimo

cionar u'_i a dichas aristas, tampoco se origina algún ciclo, entonces

$$l(u_i) = l(u'_i) \tag{2-16}$$

y, por lo tanto, P' tiene largo mínimo al igual que P. Sin embargo, P' tiene una arista común más que P con el árbol económico Q. Repitiendo varias veces esta operación obtendremos el árbol de largo mínimo que coincide con Q. Por consiguiente, Q es el árbol de largo mínimo.

2-3. REDES DE TRANSPORTE

a) Conceptos fundamentales

Se denomina *red de transporte* el grafo finito sin lazos en el cual:

1) existe un y sólo un vértice x_0 tal que $\Gamma^{-1}x_0 = \varnothing$ (este vértice se llama *entrada de la red*);

2) existe un y sólo un vértice z tal que $\Gamma z = \varnothing$ (este vértice se denomina *salida de la red*);

3) a cada arco del grafo u se refiere un número entero $c(u)$ llamado *capacidad de paso del arco u*.

El concepto de flujo está muy vinculado al de la red de transporte. Sea x un vértice arbitrario. Designemos por U_x^- el conjunto de los arcos que entran en x, y por U_x^+ el conjunto de los que

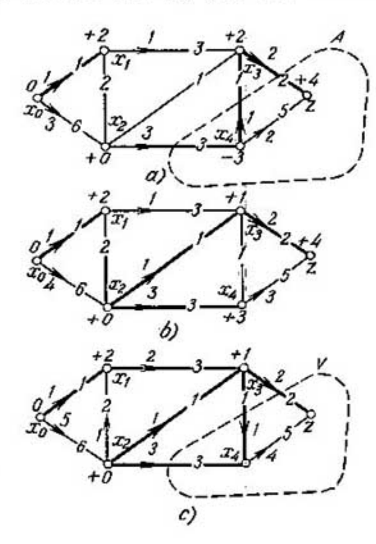


Fig. 2-8. Distribución de flujos en la red de transporte

salen de x. Se denomina flujo por la red de transporte la función $\varphi(u)$ que satisface las condiciones

$$0 \leqslant \varphi(u) \leqslant c(u), \quad u \in U; \tag{2-17}$$

$$\sum_{u \in U_x^-} \varphi(u) - \sum_{u \in U_x^+} \varphi(u) = 0; \quad x \neq x_0, \quad x \neq z. \tag{2-18}$$

La función $\varphi(u)$ puede considerarse como la cantidad de materia que pasa (en la unidad de tiempo) por el arco $u = (x, y)$ de x a y. Según la condición (2-17) esta cantidad no puede exceder de la capacidad de paso del arco $c(u)$. De acuerdo con la condición (2-18) en cada vértice x distinto de la entrada x_0 y de la salida z, la cantidad de materia que llega es igual a la que sale. Por lo tanto, ésta no puede acumularse en ningún vértice de la red de transporte excepto la entrada y salida. Esto significa que el flujo saliente del vértice de entrada x_0 es exactamente igual al

entrante en el vértice de salida z:

$$\sum_{u \in U_{x_0}^+} \varphi(u) = \sum_{u \in U_z^-} \varphi(u) = \varphi_z. \qquad (2\text{-}19)$$

La magnitud $\varphi(z)$ se denomina magnitud del flujo de la red de transporte.

En la figura 2-8 se ofrece un ejemplo de la red de transporte. Las cifras en las discontinuidades de los arcos señalan la capacidad de paso del arco. Las flechas indican la dirección de los flujos y las cifras junto a las flechas, la magnitud del mismo. Muchos problemas que surgen durante la planificación de suministros, distribución de mercancías entre consumidores, etc., se reducen al análisis de redes de transporte.

Para investigar la distribución del flujo en la red de transporte conviene introducir el concepto de corte de dicha red. Sea $A \subset X$ cierto conjunto que satisface las condiciones

$$x_0 \notin A, \quad z \in A. \qquad (2\text{-}20)$$

Designemos, respectivamente, por U_A^- y U_A^+ los conjuntos de arcos que entran en A y salen de A. Llamaremos *corte* A de la red de transporte al total de los arcos $U_A = U_A^- \cup U_A^+$. En la figura 2-8 se muestra un ejemplo de corte.

Por cuanto cada partícula de materia que se mueve de x_0 a z pasa obligatoriamente por algún arco del corte, el flujo total a través de éste será igual a la magnitud del flujo de la red de transporte, es decir, para cualquier corte A, se cumple la relación

$$\varphi_z = \sum_{u \in U_A^-} \varphi(u) - \sum_{u \in U_A^+} \varphi(u). \qquad (2\text{-}21)$$

Denominaremos capacidad de paso del corte A a la suma de las capacidades de paso de los arcos que entran en este corte:

$$c(A) = \sum_{u \in U_A^-} c(u). \qquad (2\text{-}22)$$

Ya que en cualquier arco ocurre que $\varphi(u) \leqslant c(u)$, de (2-21) y (2-22) se deduce que

$$\varphi_z \leqslant c(A). \qquad (2\text{-}23)$$

b) Problema del flujo máximo

El problema del flujo máximo en una red de transporte se enuncia del modo siguiente. Estando dada la configuración de la red de transporte y siendo conocida la capacidad de paso de los arcos, hallar la magnitud máxima del flujo que puede dejar pasar dicha red así como la distribución de este flujo en los arcos de la red.

Lema. *Si para cierta magnitud del flujo de la red de transporte* φ_z *y cierto corte* V *se cumple que* $\varphi_z = c(V)$, *el flujo* φ_z *es el mayor y el corte* V *tiene la menor capacidad de paso.*

Demostración. Como se mostró, la magnitud del flujo φ_z para cualquier corte A debe satisfacer la relación (2-23). Designemos por V el corte con la capacidad de paso mínima:

$$c(V) = \min_A c(A). \tag{2-24}$$

Como la magnitud del flujo φ_z es la misma para cualquier corte de la red de transporte, el aumento de la magnitud de dicho flujo sólo es posible hasta que éste alcance el valor $c(V)$. Por consiguiente, la magnitud del flujo

$$\varphi_z = c(V) \tag{2-25}$$

determina el flujo máximo de la red de transporte.

Sin embargo, el lema dado no ofrece todavía un método para determinar en la práctica el flujo máximo. Para formular tal método, introduzcamos algunas definiciones auxiliares.

Llamaremos *saturado* al arco u, cuando $\varphi(u) = c(u)$. Denominaremos *completo* el flujo φ_z, si cada camino de x_0 a z contiene por lo menos un arco saturado.

Para una red de transporte dada, el flujo completo no es una magnitud rigurosamente determinada y depende de la dirección de los flujos en los diferentes arcos. Así pues, en la figura 2-8, *a* y *c* se brindan dos distribuciones distintas del flujo por la misma red de transporte. Los arcos saturados están trazados con líneas gruesas. En ambos casos los flujos son completos aunque sus magnitudes son diferentes.

El algoritmo para hallar el flujo máximo, propuesto por Ford y Falkerson, consiste en el aumento paulatino del flujo φ_z hasta que éste llegue a ser máximo. En dicho caso se supone que las capacidades de paso de los arcos $c(u)$ son números enteros, de modo que los flujos en los arcos van a expresarse también con números enteros. La búsqueda del flujo completo se realiza en dos etapas.

1. *Búsqueda del flujo completo.* Sea $\varphi(u)$ cierta distribución del flujo en los arcos de la red de transporte. Busquemos el camino μ de x_0 a z todos los arcos del cual no están saturados y supongamos que

$$\varphi'(u) = \begin{cases} \varphi(u) + 1 & \text{para } u \in \mu; \\ \varphi(u) & \text{para } u \notin \mu. \end{cases} \tag{2-26}$$

Entonces el flujo φ_z cambia hasta la magnitud $\varphi'_z = \varphi_z + 1 > \varphi_z$. De este modo efectuamos el aumento paulatino de φ_z hasta que éste no se vuelva completo.

Ejemplo 2-3. Hallemos el flujo completo en la red de transporte (véase la fig. 2-8, *a*). Examinemos consecutivamente los caminos siguientes, marcando con líneas gruesas los arcos saturados:

$\mu_1 = (x_0, x_1, x_3, z)$, $\varphi(\mu_1) = 1$, se satura el arco (x_0, x_1);
$\mu_2 = (x_0, x_2, x_4, x_3, z)$, $\varphi(\mu_2) = 1$, se saturan los arcos (x_4, x_3) y (x_3, z);
$\mu_3 = (x_0, x_2, x_4, z)$; para saturar este camino puede tomarse $\varphi(\mu_3) = 2$ y se satura el arco (x_2, x_4).

Es fácil ver que no hay más caminos de x_0 a z que contengan arcos no saturados. Por lo tanto, el flujo completo es

$$\varphi_z = \varphi(\mu_1) + \varphi(\mu_2) + \varphi(\mu_3) = 4.$$

2. *Búsqueda del flujo máximo.* Sea φ_z el flujo completo y $\varphi(x, y)$, el flujo en el arco $u = (x, y)$, dirigido del vértice x al vértice y. El proceso de incremento de φ_z consiste en marcar los vértices del grafo con índices que señalen el camino en el cual es posible aumentar el flujo. Todos los vértices del grafo deben ser numerados previamente.

Marcamos x_0 con el índice 0. Si x_i es un vértice ya numerado, marcamos con el índice $+i$ todos los vértices no marcados, en los cuales van saliendo de x_i los arcos no saturados, o sea, los vértices y para los que

$$(x_i, y) \in U \text{ y } \varphi(x_i, y) < c(x_i, y), \tag{2-27}$$

y con el índice $-i$ todos los vértices no marcados de los cuales los arcos van al vértice x_i, es decir, los vértices y para los cuales

$$(y, x_i) \in U \text{ y } \varphi(y, x_i) > 0. \tag{2-28}$$

Si en este proceso resulta marcado el vértice z, entre los vértices x_0 y z habrá un circuito, todos los vértices del cual son distintos y (con una exactitud de un signo) están marcados con los números de los vértices precedentes. Aumentamos el flujo de todas las aristas de este circuito en una unidad en el sentido de x_0 a z, es decir, suponemos que

$$\varphi'(u) = \varphi(u), \qquad \text{si } u \notin \mu;$$
$$\varphi'(u) = \varphi(u) + 1, \quad \text{si } u \in \mu$$

y al moverse de x_0 a z, el arco u pasa en la dirección de su orientación;

$$\varphi'(u) = \varphi(u) - 1, \quad \text{si } u \in \mu$$

y al moverse de x_0 a z el arco u pasa en la dirección opuesta a su orientación.

Como resultado de este proceso se obtiene un nuevo flujo por la red $\varphi'_z = \varphi_z + 1$, o sea, crece la magnitud del flujo. Después se repite el proceso.

Si es imposible aumentar cierto flujo φ_z^0 por el método descrito, es decir, si resulta imposible marcar el vértice z, entonces φ_z^0 constituye el flujo máximo de la red. En efecto, sea V el conjunto de

los vértices no marcados que incluyen también el vértice z. Por lo tanto, V es un corte y precisamente aquel que no tiene arcos de salida (en caso contrario algunos vértices de este corte estarían marcados con índices negativos), y todos los arcos de entrada, saturados:

$$\left.\begin{array}{l} U_V^- = U_V; \quad U_V^+ = \varnothing; \\ \varphi(u) = c(u) \ \text{para} \ u \in U_V^-. \end{array}\right\} \qquad (2\text{-}29)$$

Siendo

$$\varphi_z^0 = \sum_{u \in U_V^-} \varphi(u) - \sum_{u \in U_V^+} \varphi(u) = \sum_{u \in U_V^-} c(u) - 0 = c(V). \quad (2\text{-}30)$$

Según el lema demostrado anteriormente φ_z^0 es el flujo mayor, y V el corte con la menor capacidad de tráfico.

Ejemplo 2-4. En la figura 2-8, *a* se muestra cómo se han marcado los índices de los vértices de la red de transporte. El vértice x resultó marcado con el índice +4. La secuencia de los índices +4, —3, +2, +0, determina el circuito $\mu = (x_0, x_2, x_3, x_4, z)$, cuyo flujo de x_0 a z hay que aumentar en una unidad. Esto conduce a la distribución del flujo mostrada en la figura 2-8, *b*. Repitiendo el proceso de marcado de los índices en este dibujo, hallamos el circuito $\mu = (x_0, x_2, x_1, x_3, x_4, z)$ cuyo flujo también debe aumentarse en una unidad. La distribución resultante del flujo se muestra en la figura 2-8, *c*. Al marcar los índices en este dibujo quedan sin marcar los vértices x_4 y z.

Por tanto, la distribución del flujo obtenida garantiza el flujo máximo φ_z^0 en la red de transporte considerada, y el conjunto $V = \{x_4, z\}$ determina el corte con la capacidad de tráfico mínima. La magnitud del flujo φ_z^0 se halla determinando la capacidad de tráfico del corte V:

$$\varphi_z^0 = c(x_2, x_4) + c(x_3, x_4) + c(x_3, z) = 3 + 1 + 2 = 6.$$

Se puede aumentar el flujo máximo de la red de transporte aumentando la capacidad de tráfico de cualquiera de los arcos incluidos en el corte V.

c) Problema de transporte

A la par con el problema de la búsqueda del flujo máximo, tiene gran importancia práctica el problema de la distribución más económica del flujo en los arcos de la red de transporte que recibió la denominación de problema de trasporte. Por cuanto en muchos casos la red de transporte constituye el esquema de organización de traslados de algunas cargas, la resolución del problema de transporte permite determinar el plan de traslados más racional, es decir, la distribución de las rutas que garantice, por ejemplo, el costo mínimo de los traslados o de la entrega de las cargas al consumidor en un tiempo mínimo. El primer problema recibió el nombre de problema de transporte según el criterio del costo y el segundo, de problema de transporte según el criterio del tiempo.

Para facilitar la exposición ulterior, designemos: $c_{ij} = c(x_i, x_j)$, la capacidad de tráfico del arco (x_i, x_j) y $d_{ij} = d(x_i, x_j)$, el costo de paso de la unidad de flujo por el arco (x_i, x_j).

El problema de transporte según el criterio del costo se puede enunciar en términos de la teoría de los grafos del modo siguiente.

Están dados la red de transporte con el flujo máximo φ_z^0 y el flujo $\varphi_z \leqslant \varphi_z^0$ que debe pasar por esta red de transporte. Se requiere hallar la distribución del flujo φ_z por los arcos de la red de transporte que asegure el costo mínimo del paso del flujo. Con esto, para cada arco debe cumplirse la relación $\varphi(x_i, x_j) \leqslant c_{ij}$, y el costo del paso del flujo $\varphi(x_i, x_j)$ por el arco (x_i, x_j) debe ser igual a $d_{ij}\varphi(x_i, x_j)$.

Para resolver este problema vamos a considerar las magnitudes d_{ij} como largos de los arcos respectivos. En este caso el costo del paso del flujo φ por cierto camino μ desde x_0 hasta z será igual al producto del largo de dicho camino por la magnitud del flujo φ y el problema de la minimización del costo del paso del flujo se reduce a la solución del problema anteriormente estudiado de la búsqueda del camino mínimo en el grafo de x_0 a z. En el caso en que no haya limitaciones en cuanto a la capacidad de tráfico de los arcos, el camino mínimo es el que garantiza el costo mínimo del paso del flujo.

Al existir limitaciones de capacidad de tráfico de los arcos el problema se resuelve en varias etapas hallando los flujos parciales en cada etapa. El método general de resolución del problema consiste en lo siguiente.

En el grafo $G_1 = (X, \Gamma)$ que representa la red de transporte con largos de arcos $d_{ij} = l(x_i, x_j)$ se halla el camino mínimo μ_1 de x_0 a z. Sea c_1 la capacidad de tráfico del camino μ_1. Por este camino se conduce el flujo

$$\varphi_l = \begin{cases} \varphi_z, & \text{si } \varphi_z \leqslant c_1; \\ c_1, & \text{si } \varphi_z > c_1. \end{cases} \quad (2\text{-}31)$$

Si $\varphi_z \leqslant c_1$, el problema está resuelto y μ_1 es el camino más económico para el flujo φ_z.

Si $\varphi_z > c_1$, entonces consideramos a φ_1 como flujo parcial y pasamos al grafo G_2 que se obtiene del gráfo G_1 sustituyendo las capacidades de tráfico de los arcos c_{ij} por c'_{ij} de la relación

$$c'_{ij} = \begin{cases} c_{ij} - c_1 & \text{para } u \in \mu_1; \\ c_{ij} & \text{para } u \notin \mu_1. \end{cases} \quad (2\text{-}32)$$

Por cierto, se excluyen del análisis los arcos en los cuales $c'_{ij} = 0$. El flujo cuya distribución se busca en el grafo G_2 se toma igual a

$$\varphi'_z = \varphi_z - \varphi_1. \quad (2\text{-}33)$$

Ahora surge el problema inicial de buscar la distribución más económica del flujo φ'_z pero ya con respecto al grafo G_2. Su resolución da el camino μ_2 con la capacidad de tráfico c_2 a través del

cual se hace pasar el flujo parcial

$$\varphi_2=\begin{cases}\varphi'_z, & \text{si } \varphi'_z \leqslant c_2;\\ c_2, & \text{si } \varphi'_z > c_2.\end{cases} \qquad (2\text{-}34)$$

Si $\varphi'_z \leqslant c_2$ el problema está resuelto y la distribución de flujos más económica en el grafo G_1 será el paso del flujo φ_1 por el camino μ_1 y del flujo φ_2 por el camino μ_2.

Si $\varphi'_z > c_2$ hay que recurrir a un nuevo grafo G_3 y hallar un nuevo flujo parcial φ_3. Este proceso se repite hasta que la suma de los flujos parciales alcance el valor φ_z. Estos flujos parciales conducidos por el grafo G_1 proporcionan, en efecto, la distribución más económica del flujo φ_z.

Para ilustrar el método descrito, examinaremos la variante más difundida del problema de transporte según el criterio del costo.

En las estaciones $x_1, \ldots, x_m$ hay una carga semejante en cantidades $a_1, \ldots, a_m$. Se requiere conducirla en cantidades $b_1, \ldots, b_r$ a las estaciones $y_1, \ldots, y_r$. Se supone que la cantidad total de carga requerida es igual a las reservas existentes:

$$\sum_{i=1}^{m} a_i=\sum_{j=1}^{r} b_j. \qquad (2\text{-}35)$$

Tabla 2-1

Problema de transporte

a_i	b_j		
	b_1	...	b_r
a_1	d_{11}	...	d_{1r}
...	...	...	...
a_m	d_{m1}	...	d_{mr}

El costo del traslado de la carga de la estación x_i a la y_j es igual a d_{ij}. Se necesita hallar las rutas más económicas para transportar las cargas. Es cómodo anotar los datos iniciales en forma de la tabla 2-1.

La red de transporte correspondiente a este problema se construye del modo siguiente. La entrada x_0 se une a cada uno de los vértices x_i mediante un arco con capacidad de tráfico $c(x_0, x_i)=a_i$. Cada uno de los vértices y_j se une a la salida z por medio de un arco con capacidad de tráfico $c(y_j, z)=b_j$. El costo de paso del flujo por los arcos (x_0, x_i) e (y_j, z) se considera igual a cero. Finalmente, cada vértice x_i se une con cada vértice y_j mediante un arco de capacidad de tráfico infinita. El costo de paso por éste de la unidad de flujo es igual a d_{ij}. Más adelante se aplica a esta red de transporte el método ya estudiado.

Ejemplo 2-5. Hallar las rutas más económicas para el problema de transporte planteado en la tabla 2-2. Esta tabla corresponde al esquema de los caminos que comunican las fábricas productoras de piezas de construcción con los consumidores de dichas piezas (obras), mostrado en la figura 2-9. En la figura 2-10 se ofrece la red de transporte correspondiente a los datos de la tabla 2-2.

Utilizando el método descrito más arriba hallamos los flujos parciales y las rutas enumerados en la tabla 2-3 en el orden de su obtención. El plan de transporte correspondiente a esta tabla se muestra asimismo en el esquema de caminos.

No es difícil ver que en el caso general el costo de traslado de las cargas por la red de transporte del tipo considerado se determina mediante la expresión

$$\sum_{i=1}^{m} \sum_{j=1}^{r} d_{ij}\varphi(x_i, y_j). \quad (2\text{-}36)$$

Tabla 2-2

Datos iniciales del problema de transporte

a_i	b_j			
	5	10	20	15
10	8	3	5	2
15	4	1	6	7
25	1	9	4	3

Por consiguiente, el método analizado para resolver el problema de transporte, brinda en esencia la forma de hallar las magnitudes de los flujos parciales $\varphi(x_i, y_j)$ que minimizan la suma indicada. El método de resolución examinado no es único. Encontraremos problemas similares en la sección dedicada a la programación lineal donde se mostrarán otros métodos para resolver problemas análogos.

Analicemos la *resolución del problema de transporte por el criterio del tiempo* mediante el ejemplo de red de transporte dada en

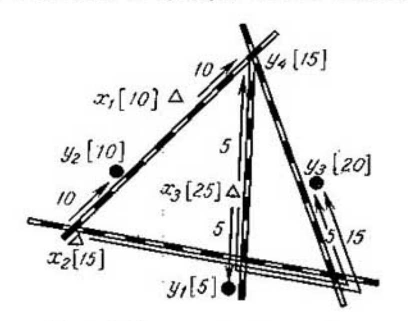


Fig. 2-9. Esquema de los caminos

la tabla 2-2 en la que interpretaremos ahora las magnitudes d_{ij} como el tiempo requerido para trasladar la carga del punto x_i al punto y_j y las designaremos por t_{ij}. Es posible confrontar problemas similares sobre el transporte de productos de fácil deterioro, el envío de medios de auxilio a regiones de cataclismos, el acarreo de granos de la nueva cosecha a los puntos de acopio, etc. En to-

Tabla 2-3

Distribución de los flujos parciales en el problema de transporte por el criterio de costo

k	Ruta (x_i, y_j)	Flujo parcial φ_k	d_{ij}	Costo del traslado $d_{ij}\varphi_k$
1	(x_3, y_1)	5	1	5
2	(x_2, y_2)	10	1	10
3	(x_1, y_4)	10	2	20
4	(x_3, y_4)	5	3	15
5	(x_3, y_3)	15	4	60
6	(x_2, y_3)	5	6	30
—	—	50	—	140

das estas tareas es necesario trasladar todas las cargas a los lugares de destino en el más breve intervalo de tiempo.

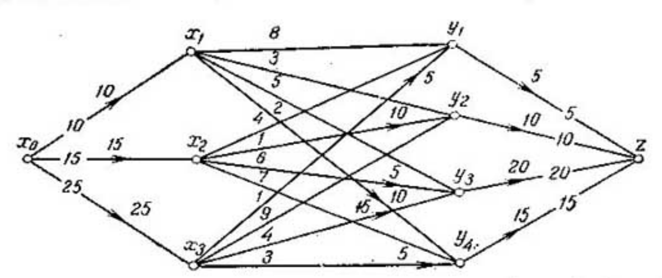


Fig. 2-10. Solución del problema de transporte por el criterio de costo

Examinemos el método general de resolución de este problema. Supongamos que de cualquier modo se ha encontrado cierta distribución del flujo φ_z en el grafo G que representa la red de transporte considerada. Separemos del grafo G el grafo parcial G' en el cual sólo incluimos los arcos que participan en la transmisión del flujo φ_z. Sea μ cierto camino que conduce de x_0 a z y t_μ, el tiempo de paso del flujo por este camino. Evidentemente, el tiempo necesario para trasladar todas las cargas de x_0 a z va a determinarse por el camino que tiene la mayor duración de paso del flujo ya que el traslado de las cargas por los demás caminos terminará antes. Por lo tanto, el tiempo T, requerido para el transporte de todas las cargas será igual a:

$$T = \max_{\mu \in G'} t_\mu. \tag{2-37}$$

La resolución del problema de transporte por el criterio del tiempo se reduce, de esta manera, a separar del grafo G, el grafo parcial G' que sea capaz de dejar pasar todo el flujo φ_z y en el cual la duración del camino más prolongado sea mínima en comparación con todos los demás grafos similares. Por cierto, la solución hallada por el criterio del costo antes descrito, que minimiza la magnitud definida por la expresión (2-36), al mismo tiempo puede ser no la mejor desde el punto de vista del criterio del tiempo.

La solución de la tarea planteada se reduce al mejoramiento consecutivo del grafo G', eliminando del mismo los caminos más prolongados e introduciendo los más cortos pero que no se han utilizado antes, y redistribuyendo correspondientemente el flujo φ_z.

Sirvámonos del ejemplo examinado. Tomemos a modo de primera aproximación a la solución óptima, la solución obtenida basándose en el criterio del

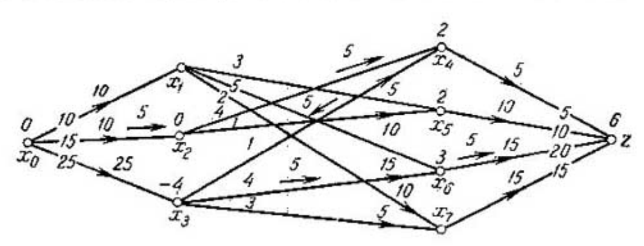


Fig. 2-11. Solución del problema de transporte por el criterio de tiempo

costo. La distribución de los flujos para este caso se muestra en la figura 2-10. En esta figura vemos que el tiempo de paso del flujo por la ruta más prolongada (x_2, y_3) es igual a 6.

Sin embargo, quedaron sin utilizar las rutas menos largas (x_1, y_2), (x_2, y_1), (x_1, y_3). Por eso puede ser que el empleo de cualquiera de estas rutas permita excluir la ruta (x_2, y_3).

Construyamos el grafo parcial G' en el cual sólo incluimos los arcos del grafo G que tienen $t_{ij} < 6$ y en el cual quedó la misma distribución del flujo que en el grafo G. El grafo G' se muestra en la figura 2-11 donde, por conveniencia, los vértices y_j están designados por $x_{m+j} = x_{3+j}$. El flujo φ'_z a través de este grafo es igual a 45 unidades, o sea, menos que el flujo inicial $\varphi_z =$ $= 50$ unidades. Este flujo es completo ya que en el grafo G' todos los caminos de x_0 a z tienen arcos saturados. Sin embargo, puede que éste no sea el máximo.

Si el flujo máximo en el grafo G' es igual a φ_z, entonces se hallará una distribución de este flujo, con la cual el camino más prolongado va a durar un tiempo $t_\mu < 6$ unidades. Por tanto, la solución ulterior del problema se reduce a determinar el flujo máximo en el grafo G'.

En la figura 2-11 se ha efectuado la marcación de los vértices del grafo G' en concordancia con la regla del § 2-3. El marcado de los vértices indica la existencia de un camino $(x_0, x_2, x_4, x_3, x_6, z)$, en el cual el flujo en la dirección de x_0 a z puede incrementarse en 5 unidades. Este flujo adicional se muestra con flechas. Como vemos, el flujo máximo en el grafo G' es $\varphi_z = 50$ unidades y la ruta más prolongada tiene un tiempo de 4 unidades. No es posible

Tabla 2-4

Distribución de los flujos parciales en el problema de transporte por el criterio de tiempo

Ruta	Flujo parcial	Tiempo de paso del flujo
(x_2, y_2)	10	1
(x_1, y_4)	10	2
(x_3, y_4)	5	3
(x_2, y_1)	5	4
(x_3, y_3)	20	4
—	$\varphi_z = 50$	$t_{máx} = 4$

lograr la disminución ulterior del tiempo de paso del flujo, ya que al vértice $y_8(x_6)$ no van rutas con el tiempo menor de 4 unidades.

En la tabla 2-4 se presenta la distribución definitiva de los flujos parciales por las rutas que da la solución del problema de transporte según el criterio del tiempo.

Capítulo tercero

ESPACIOS MULTIDIMENSIONALES

3-1. ESPACIOS METRICOS Y DISTANCIAS

a) Concepto de distancia

El estudio de los conjuntos realizado en el capitulo 1 como reunión de ciertos objetos tiene una aplicación limitada, ya que en la naturaleza todos los objetos materiales están en interrelación e interacción. Por eso, es necesario vincular el concepto de conjunto con el establecimiento de unas u otras correlaciones entre sus elementos.

Se dice que un conjunto tiene *estructura* si entre los elementos del conjunto se han establecido determinadas relaciones o con los mismos se han definido ciertas operaciones. El conjunto que posee estructura se denomina *espacio.*

Comencemos el estudio de los espacios por el tipo de espacios más sencillos llamados espacios métricos, para cuya definición es necesario introducir el concepto de distancia entre los elementos del conjunto.

El hombre tropieza cada dia con el concepto de distancia relacionando este concepto con la distribución espacial de los objetos e interpretando por distancia la medida del alejamiento de los objetos, unos de otros. Generalmente la distancia $d(M, N)$ entre los puntos M y N se mide por el largo del segmento que une dichos puntos (fig. 3-1). Sin embargo, tal distribución de la distancia con

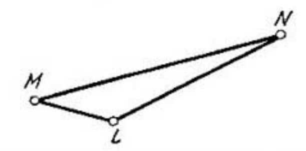


Fig. 3-1. Representación del axioma del triángulo

frecuencia resulta insuficiente. Así, incluso en usanza, la distancia entre dos ciudades no se determina unívocamente (distancia por ferrocarril, vía marítima, fluvial, etc.). En una ciudad dividida en barrios como se muestra en la figura 3-2, no tiene sentido medir la distancia por medio del segmento de recta que une los puntos M y N ya que sólo es posible trasladarse por las calles.

Por otro lado, frecuentemente no relacionamos la palabra alejamiento con el espacio en su sentido habitual, sino que nos refe-

rimos al alejamiento en el tiempo (en la profunda lejanía de los siglos, etc.) o en otro sentido. Si cierto sistema puede tomar sucesivamente los estados $A_1, A_2, \ldots, A_n$, entonces podemos medir el alejamiento del estado A_k respecto a A_j con el número de estados por los cuales debe atravesar el sistema para pasar del estado A_j al A_k. Aquí el estado va a ser la medida del alejamiento, uno del otro, de los estados del sistema. Pero si se considera el alejamiento como propiedad del estado, nos vemos obligados a hablar no ya del espacio tridimensional habitual sino de algún otro espacio que se puede denominar, por ejemplo, espacio de estados.

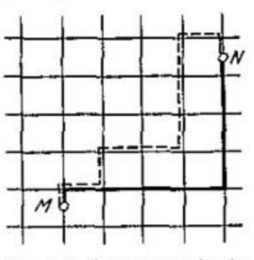


Fig. 3-2. Determinación de la distancia en una ciudad dividida en barrios

En los ejemplos dados se ve que debe existir cierta definición general de la distancia como medida del alejamiento de los objetos y, por consiguiente, también del espacio en que existen estos objetos, con esto, en distintas situaciones concretas dichos conceptos pueden tener diferente contenido. Por cuanto las reuniones de diversos objetos constituyen conjuntos, los conceptos de espacio y distancia deben estar vinculados al concepto de conjunto.

b) Definición de espacio métrico

Sea X un conjunto arbitrario. El concepto de distancia entre los elementos de X se obtiene mediante la generalización de las propiedades fundamentales, que pueden esperarse intuitivamente del concepto de distancia y que pueden entenderse con facilidad al examinar la figura 3-1.

Relacionemos con cada par de elementos de X cierto número real no negativo $d \geqslant 0$. Este número se llama *distancia* o *métrica* en X si para cualesquiera $x, y, z \in X$ el mismo satisface las tres condiciones siguientes:

1) $d(x, y) = 0$ cuando, y sólo cuando $x = y$ (axioma de identidad);

2) $d(x, y) = d(y, x)$ (axioma de simetría);

3) para cualquier tríada $x, y, z \in X$ se cumple $d(x, y) \leqslant \leqslant d(x, z) + d(z, y)$ (axioma del triángulo).

Se llama *espacio métrico* al par (X, d), es decir, al conjunto X con la métrica d definida en éste. Los elementos del conjunto X se denominan puntos del espacio métrico (X, d).

De la definición dada se deduce que el conjunto X sólo se transforma en espacio métrico cuando se ha introducido en él la métrica correspondiente $d(x, y)$. Si en un mismo conjunto X se introducen distintas métricas, se obtienen asimismo distintos espa-

cios. Así pues, los espacios representados en las figuras 3-1 y 3-2, tienen en calidad de elementos conjunto de puntos de un plano, pero poseen distintas métricas.

c) Ejemplos de espacios métricos

1. Sean x, y elementos cualesquiera del conjunto R de los números reales. El conjunto R puede transformarse en espacio métrico si se determina la distancia entre x e y por la fórmula

$$d(x, y) = |x - y|. \qquad (3\text{-}1)$$

Precisamente por esta fórmula se halla la distancia entre los puntos del eje real que constituye uno de los más sencillos ejemplos del espacio métrico.

2) Examinemos el conjunto R^n cuyos elementos son las n-es ordenadas de los números reales del tipo $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$ ... El conjunto R^n puede transformarse en espacio métrico por distintos métodos.

La distancia entre los puntos x e y se define muy frecuentemente por la fórmula

$$d_2(x, y) = \left\{ \sum_{i=1}^{n} |x_i - y_i|^2 \right\}^{1/2}. \qquad (3\text{-}2)$$

En el caso de $n = 2, 3$ esta definición coincide con el concepto habitual de distancia. Las propiedades 1, 2 y 3 para esta distancia son evidentes al examinar la figura 3-1.

La métrica $d_2(x, y)$ se denomina euclídea y el espacio R^n con tal métrica se llama euclídeo y se designa por E_n.

3. Para el conjunto R^n la distancia puede determinarse por otros métodos, por ejemplo así:

$$d_1(x, y) = \sum_{i=1}^{n} |x_i - y_i| \qquad (3\text{-}3)$$

o bien

$$d_\infty(x, y) = \text{máx}\,(|x_1 - y_1|, \ldots, |x_n - y_n|). \qquad (3\text{-}4)$$

Es fácil ver que las métricas d_2, d_1, d_∞ son casos particulares de la métrica

$$d_p(x, y) = \left\{ \sum_{i=1}^{n} |x_i - y_i|^p \right\}^{1/p}$$

y se obtienen, respectivamente, con $p = 2$, $p = 1$ y $p = \infty$.

Las propiedades 1 y 2 son evidentes para las métricas (3-3) y (3-4). Para demostrar la propiedad 3 introduzcamos un punto

más $z=(z_1, \ldots, z_n) \in R^n$. Para la distancia $d_1(x, y)$ tenemos:

$$d_1(x, y)=\sum_{i=1}^{n} |x_i - z_i + z_i - y_i| \leqslant$$
$$\leqslant \sum_{i=1}^{n} (|x_i - z_i| + |z_i - y_i|) = d(x, z) + d(z, y).$$

Para la distancia $d_\infty(x, y)$ la propiedad 3 se comprueba del modo siguiente. Supongamos que $|x_k - y_k|$ es la mayor de las diferencias correspondientes de los puntos x e y. Entonces

$$d_\infty(x, y)=|x_k - y_k| = |x_k - z_k + z_k - y_k| \leqslant |x_k - z_k| + |z_k - y_k|.$$

Es evidente que

$$|x_k - z_k| \leqslant \text{máx}(|x_1 - z_1|, \ldots, |x_n - z_n|) = d_\infty(x, z);$$
$$|z_k - y_k| \leqslant \text{máx}(|z_1 - y_1|, \ldots, |z_n - y_n|) = d_\infty(z, y).$$

Luego,

$$d_\infty(x, y) \leqslant d_\infty(x, z) + d_\infty(z, y).$$

4. Examinemos un conjunto de funciones del tiempo de todo género posible continuas en el intervalo $a \leqslant t \leqslant b$. Sean $x(t)$ e $y(t)$ dos de estas funciones. La distancia entre ellas puede determinarse mediante la relación

$$d(x, y) = \max_{a \leqslant t \leqslant b} |x(t) - y(t)|, \tag{3-5}$$

que, como es fácil comprobar, satisface todas las propiedades de la métrica. El espacio con esta métrica se designa por $C_{[a, b]}$.

Para ilustrar las posibilidades de utilizar en la práctica los conceptos introducidos conviene detenerse en un espacio sumamente importante en cibernética que se denomina espacio de mensajes.

3-2. INTERPRETACION GEOMETRICA DE LAS SEÑALES Y MENSAJES

a) Espacio de mensajes

Como hemos visto en la Introducción, los mensajes se transmiten por canales de comunicación con ayuda de cierto alfabeto $\mathfrak{A}$, que consta de un número finito de símbolos. Los mensajes constituyen distintas secuencias de símbolos del alfabeto. El número de símbolos en un mensaje se llama largo del mensaje.

Analicemos el caso en que con ayuda del alfabeto $\mathfrak{A}_m$ que consta de m símbolos se transmiten mensajes de largo n. Pueden catalogarse también aquí los mensajes más breves si éstos se complementan hasta el largo n con cierto símbolo determinado, por ejemplo, con el cero en caso de utilizarse el alfabeto binario. El

conjunto de todos estos mensajes puede considerarse como espacio métrico si se introduce el concepto de distancia entre mensajes.

Llamemos distancia $d(x, y)$ entre dos mensajes x e y al número de posiciones en las cuales los mensajes x e y tienen diferentes símbolos. El espacio métrico obtenido en este caso se designará por $E(n, \mathfrak{A}_m)$ y se llamará *espacio de mensajes.*

Ejemplo 3-1. $\mathfrak{A}_m$ es el alfabeto ruso, $n = 7$, $x =$ (картина), $y =$ (корзина). No coinciden la segunda y cuarta letras. Luego, $d(x, y) = 2$.

Ejemplo 3-2. Sea $\mathfrak{A} = \mathfrak{A}_2 = \{0, 1\}$ el alfabeto binario, $n = 10$, $x = 0100111010$, $y = 0010110010$. No coinciden el segundo, tercero y séptimo signos. Por tanto, $d(x, y) = 3$.

La definición dada de la distancia $d(x, y)$, en el espacio de mensajes, satisface todas las propiedades de la distancia. El axioma de identidad y el de simetría resultan evidentes. Cerciarémonos de que también se cumple el axioma del triángulo.

Examinemos tres mensajes x, y, z de largo n. Designemos por x_k, y_k y z_k los símbolos de k-ésima posición de estos mensajes. Naturalmente, si $x_k = y_k$ e $y_k = z_k$, entonces $x_k = z_k$, es decir, si en alguna posición coinciden los símbolos de los mensajes x, y, y de los mensajes y, z, entonces en dicha posición coinciden también los símbolos de los mensajes x y z. Por lo tanto, en los mensajes x y z solamente pueden haber símbolos que no coinciden en aquellos lugares donde no coinciden los símbolos bien en los mensajes x, y, bien en los mensajes y, z. Pero esto significa que el número total de símbolos no coincidentes en los mensajes x y z no puede sobrepasar la suma de números de los símbolos no coincidentes en los mensajes x, y e y, z.

En lo sucesivo, con fines de simplicidad y claridad, nos limitaremos únicamente al análisis del alfabeto binario $\mathfrak{A} = \mathfrak{A}_2 = \{0, 1\}$. Esto puede hacerse sin detrimento de la generalidad de los razonamientos ya que los mensajes representados por símbolos de un alfabeto siempre pueden recodificarse como mensajes representados mediante símbolos de otro alfabeto. Examinemos, por ejemplo, el alfabeto $\mathfrak{A}_m$ que contiene m símbolos. Yuxtapongamos a cada símbolo un número de orden de 0 a $m - 1$. Si se sustituyen en el mensaje los símbolos del alfabeto $\mathfrak{A}_m$ por sus números de orden en el sistema de numeración binario, obtenemos la representación del mensaje en el alfabeto binario.

b) Concepto de códigos estables a las interferencias

En el proceso de transmisión por un canal de comunicación puede distorsionarse el mensaje. En el caso del alfabeto binario la distorsión consiste en que en el mensaje captado algunas unidades resultan cambiadas por ceros y algunos ceros por unidades. Surge la cuestión de si pueden elaborarse códigos que permitan descubrir que en el proceso de la transmisión ocurrió una distorsión o hasta restablecer los valores de los dígitos distorsionados.

Los códigos que poseen estas propiedades se denominan *estables a las interferencias.*

Examinemos el caso en que durante el proceso de transmisión del mensaje no puede ocurrir una distorsión mayor que en k dígitos. En el espacio de los mensajes $E(n, \mathfrak{A}_2)$ separamos el subconjunto $H_k \subseteq E(n, \mathfrak{A}_2)$ que posee la propiedad de que para $x, y \in H_k$ cualesquiera, se cumple:

$$d(x, y) > k. \tag{3-6}$$

Llamemos al conjunto H_k conjunto de palabras comprensibles. Aquí el término "palabra" es sinónimo del término "mensaje". Entonces cualquier $x \notin H_k$ es una palabra incomprensible. Supongamos que al transmitir la palabra $x \in H_k$ ésta se distorsionó y convirtió en la palabra x. Como que por hipótesis no pueden ocurrir más de k distorsiones, entonces $d(x, x') \leqslant k$ y $x' \notin H_k$, o sea, x' es una palabra incomprensible. Así pues, la recepción de una palabra incomprensible denota que ocurrió una distorsión en el proceso de transmisión. Los códigos que satisfacen la condición (3-6) se denominan códigos con detección de error.

Ejemplo 3-3. Separemos del código $E(3, \mathscr{A}_2)$ el siguiente conjunto de palabras comprensibles que satisfacen la condición $d(x, y) = 2$:

$$H_1 = \{000, 101, 011, 110\}.$$

La distorsión de cualquiera de los dígitos en estas palabras las convierte en incomprensibles, o sea, permite detectar un error sencillo.

El conjunto H_1 que forma un código con detección de error sencillo en palabras de largo n puede obtenerse del modo siguiente. Examinemos el conjunto $E(n-1, \mathfrak{A}_2)$, es decir, el conjunto de palabras de largo $n-1$. El conjunto H_1 se obtiene añadiendo a dicho conjunto un dígito más cuyo valor se elige de modo que sea par el número total de unidades en las palabras $x \in H_1$.

Ejemplo 3-4. Para $n = 4$ tenemos:

$$E(3, \mathfrak{A}_2) = \{000, 001, 010, 100, 011, 101, 110, 111\}.$$

Por cierto

$$H_1 = \{0000, 0011, 0101, 1001, 0110, 1010, 1100, 1111\}.$$

Es cómodo descubrir la palabra distorsionada con ayuda de la operación de adición por el módulo 2 según las reglas:

$$0 \oplus 0 = 0;\ 0 \oplus 1 = 1;\ 1 \oplus 0 = 1;\ 1 \oplus 1 = 0. \tag{3-7}$$

Así pues, en la palabra $x = a_1 a_2 \ldots a_n$ la magnitud

$$\beta = a_1 \oplus a_2 \oplus \ldots \oplus a_n$$

será cero o unidad dependiendo de si tienen el valor de la unidad un número par o impar de símbolos de la palabra.

Observemos que el código de cuatro dígitos con detección de error que se obtuvo en el ejemplo 3-4, brinda la posibilidad de componer $2^4 = 16$ palabras diferentes, aunque en calidad de palabras comprensibles, o sea, destinadas a transmitirse por el canal de comunicación, se utilizan solamente 8. Los códigos en los que el número de palabras comprensibles es menor que el total de palabras posibles, recibieron la denominación de *códigos con re-*

dundancia. La existencia de redundancia es condición necesaria para la elaboración de códigos estables a las interferencias.

Resulta igualmente posible elaborar códigos que permitan corregir los errores cometidos. Supongamos nuevamente que en el proceso de transmisión puedan distorsionarse no más de k dígitos del código. El conjunto de palabras comprensibles $H_k = E(n, \mathfrak{A}_2)$ se elige por la condición

$$d(x, y) > 2k \qquad (3\text{-}8)$$

para $x, y \in H_h$ cualesquiera.

Consideremos dos palabras cualesquiera $x, y \in H_k$. Supongamos que como resultado de una distorsión x se convirtió en x'. Entonces, $d(x, x') \leqslant k$. Por la desigualdad del triángulo obtenemos:

$$d(x', y) \geqslant d(x, y) - d(x, x') > 2k - k = k. \qquad (3\text{-}9)$$

Consecuentemente,

$$d(x, x') < d(x', y). \qquad (3.10)$$

De esta manera, la distancia desde la palabra equivocada x' hasta la x que ha sido sometida a distorsión, es menor que hasta cualquier otra palabra comprensible. Hallando la palabra comprensible más cercana a x', nosotros, de este modo, restablecemos el mensaje concreto x. Los códigos que satisfacen la condición (3-8) se denominan *códigos con corrección de errores*. Las cuestiones de la realización práctica de los códigos con corrección de errores son bastante complejas y se estudian en literatura especializada.

3-3. ESPACIOS LINEALES NORMADOS

a) Espacio lineal

Al comienzo del capítulo se definió el espacio como conjunto dotado de una estructura determinada. Se estudiaron los espacios métricos cuya estructura se determinaba por el hecho de que a cada par de elementos se atribuía un número real llamado métrica que satisfacía determinadas propiedades. No obstante, la introducción de la métrica está lejos de agotar todas las propiedades estructurales de los diversos espacios. En particular, si el conjunto X consta de números reales o complejos, se incluyen entre las propiedades estructurales de importancia la posibilidad de obtener elementos del conjunto a partir de otros mediante la adición de estos elementos, o la multiplicación de un elemento por un escalar. Los conjuntos que tienen dichas propiedades pertenecen a la clase de los espacios lineales. Los espacios lineales deben satisfacer las condiciones siguientes:

1) para cada par de elementos, $x, y \in X$ está definido unívocamente un tercer elemento $z \in X$ que se denomina su suma y se designa por $x + y$, siendo que

$x + y = y + x$ (conmutatividad);

$x + (y + v) = (x + y) + v$ (asociatividad);

en X hay un elemento 0 tal que $x + 0 = x$ para todas las $x \in X$ (existencia del cero);

para cada $x \in X$ hay un elemento $-x$ tal que $x + (-x) = 0$ (existencia del elemento opuesto);

2) para cualquier número α y cualquier elemento $x \in X$ está determinado el elemento $\alpha x \in X$ siendo que

$$(\alpha + \beta) x = \alpha x + \beta x;$$

$$\alpha (x + y) = \alpha x + \alpha y.$$

Las condiciones 1) y 2) se denominan condiciones de aditividad y homogeneidad del espacio lineal.

Pueden servir de ejemplos de espacios lineales:

1. El conjunto de números reales R con la definición habitual de las operaciones de adición y multiplicación.

2. El conjunto R_n de todas las n-es ordenadas de los números reales, si las operaciones de adición y multiplicación por un número están definidas del modo siguiente. Si $x, y \in R^n$ y $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$, entonces, $x + y = (x_1 + y_1, \ldots, x_n + y_n)$; $\alpha x = (\alpha x_1, \ldots, \alpha x_n)$.

Para $n = 2$ y $n = 3$ estas operaciones coinciden con las reglas habituales de operación con vectores. Se entiende por vector nulo el vector $(0, \ldots, 0)$ con componentes iguales a cero.

3. El espacio de las funciones $C_{[a,\, b]}$, si para $x(t), y(t) \in C_{[a,\, b]}$ cualesquiera se entiende por $x(t) + y(t)$ la suma de los valores $x(t)$ e $y(t)$ tomados con los mismos valores de t, y por $\alpha x(t)$ se comprende una nueva función obtenida de $x(t)$, multiplicando todos sus valores por α. La función nula será $f(t) \equiv 0$ que es idénticamente igual a cero en todo el intervalo $[a, b]$.

b) Espacio lineal normado

El espacio lineal sólo recibe su descripción definitiva cuando las propiedades de aditividad y homogeneidad están complementadas por la posibilidad de medir la magnitud de los propios elementos. Así pues, no podemos comparar vectores si no convenimos en qué se comprende por magnitud (largo) del vector. La introducción en el espacio lineal de los estimados numéricos de las magnitudes de los distintos elementos conduce al concepto de espacio lineal normado denominado a veces espacio de Banach.

El espacio lineal se denomina espacio lineal normado, si para cada $x \in X$ existe un número no negativo $\|x\|$ llamado *norma x*, que satisface las condiciones siguientes:

$\| x \| = 0$ cuando, y sólo cuando $x = 0$;
$\| \alpha x \| = |\alpha| \cdot \| x \|$;
$\| x + y \| \leqslant \| x \| + \| y \|$ (desigualdad del triángulo).
No es difícil establecer que la magnitud $\|x - y\|$ posee todas las propiedades de la distancia $d(x, y)$ en el espacio métrico. Efectivamente:
$\| x - y \| = 0$, si $x - y = 0$, es decir, si $x = y$;
teniendo en cuenta que $y - x = -(x - y)$, hallamos:

$$\| y - x \| = | -1 | \cdot \| x - y \| = \| x - y \|;$$

$$\| x - y \| = \| (x - z) + (z - y) \| \leqslant \| x - z \| + \| z - y \|.$$

Por tanto, el espacio lineal normado es un espacio métrico con métrica

$$d(x, y) = \| x - y \|. \tag{3-11}$$

Todos los espacios métricos anteriormente examinados complementados con las propiedades de aditividad y homogeneidad, se transforman en espacios lineales normados. Para dichos espacios se emplean notaciones especiales, a saber:

1) el espacio $C_2^{(n)}$ o E_n con la norma

$$\| x \| = \left\{ \sum_{i=1}^{n} | x_i |^2 \right\}^{1/2} \quad \text{o} \quad \| x \| = | x | \text{ para } n = 1; \tag{3-12}$$

2) el espacio $C_1^{(n)}$ con la norma

$$\| x \| = \sum_{i=1}^{n} | x_i |; \tag{3-13}$$

3) el espacio $C^{(n)}$ con la norma

$$\| x \| = \max \{ |x_1|, \ldots, |x_n| \}; \tag{3-14}$$

4) el espacio $C_{[a, b]}$ de las funciones continuas definidas en el intervalo $[a, b]$ con la norma

$$\| f \| = \max_{a \leqslant t \leqslant b} | f(t) |. \tag{3-15}$$

3-4. APLICACION DE LOS ESPACIOS MULTIDIMENSIONALES EN ALGUNOS PROBLEMAS DE CIBERNETICA

a) Atenuación de errores en datos experimentales

Los resultados de las observaciones de cierta magnitud física y son corrientemente una secuencia de valores medidos de esta magnitud $(x_1, \ldots, x_n) = x$, por cierto, los resultados de las mediciones habitualmente contienen errores provocados por las imperfecciones

del experimento y la influencia de diversos factores extraños. La misma magnitud y puede no permanecer constante sino variar según cierta ley. El objetivo del experimento es establecer el valor real de la magnitud observada.

La situación descrita puede representarse en términos de la teoría de la transmisión de mensajes si se considera y como cierto mensaje acerca del valor de la magnitud a medir, sobre el cual se superponen las distorsiones debidas a las imperfecciones del experimento. Representando el mensaje transmitido y y el recibido x en forma de puntos del espacio de mensajes, puede estimarse el grado de distorsión del mensaje recibido con ayuda del valor de la distancia $d(x, y)$ cuyo método de determinación depende del carácter del experimento realizado. Con mucha frecuencia se utiliza distancia del tipo

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2}, \quad (3\text{-}16)$$

con la cual el espacio de mensajes se transforma en espacio $C_2^{(n)}$.

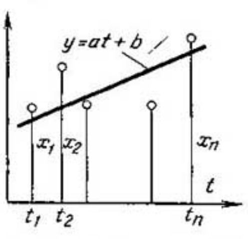


Fig. 3-3. Determinación de los parámetros de una señal que varía linealmente

Por lo general, mediante consideraciones teóricas, se logra establecer el conjunto de los mensajes teóricamente posibles Y. En este caso debe tomarse como mensaje correcto tal $y \in Y$ que sea más cercano al mensaje recibido x, es decir, para el cual

$$d(x, y) = \text{mín}, \quad (3\text{-}17)$$

o, lo que es a veces más conveniente,

$$d^2 = (x, y) = \text{mín}. \quad (3\text{-}18)$$

El principio de determinación del mensaje verdadero y por la fórmula (3-18) para la distancia del tipo (3-16) se utiliza ampliamente con el nombre de método de los cuadrados mínimos.

Examinemos a modo de ejemplo el caso de importancia práctica en que la magnitud y varía según la ley lineal:

$$y = at + b. \quad (3\text{-}19)$$

En los momentos $t_1, \dots, t_n$ se miden los valores $y_k = at_k + b$, $k = 1, \dots, n$. Los resultados de las mediciones $x_1, \dots, x_n$ están representados en la figura 3-3.

Tomando en consideración (3-19) la condición (3-18) se escribe en la forma

$$d^2(x, y) = F(a, b) = \sum_{k=1}^{n} (x_k - at_k - b)^2 = \text{min}. \quad (3\text{-}20)$$

En esta expresión son incógnitos los parámetros a y b de la ley lineal buscada de variación de la magnitud y. Igualando a cero las derivadas parciales por a y b de la función $F(a, b)$, obtenemos dos ecuaciones que contienen ambas incógnitas y que después de simplificarlas adquieren la forma:

$$\left.\begin{aligned} a\sum_{k=1}^{n} t_k^2 + b\sum_{k=1}^{n} t_k &= \sum_{k=1}^{n} t_k x_k; \\ a\sum_{k=1}^{n} t_k + nb &= \sum_{k=1}^{n} x_k. \end{aligned}\right\} \qquad (3\text{-}21)$$

No es difícil determinar en (3-21) los valores de a y b, que nos interesan.

Otro caso importante tiene lugar cuando la magnitud y se mantiene invariable, es decir,

$$y = c = \text{const} \qquad (3\text{-}22)$$

y para su estimación se han realizado n mediciones independientes $x_1, \ldots, x_n$. Teniendo en cuenta (3-16) la condición (3-18) se escribe en la forma

$$d^2(x, y) = F(c) = \sum_{k=1}^{n} (x_k - c)^2 = \text{mín}. \qquad (3\text{-}23)$$

Diferenciando esta expresión por c e igualando a cero, hallamos:

$$c = \frac{1}{n}\sum_{k=1}^{n} x_k. \qquad (3\text{-}24)$$

El valor c determinado por la fórmula (3-24) se denomina media aritmética de los valores $x_1, \ldots, x_n$.

Con una definición distinta de la distancia se obtienen otras fórmulas para determinar la media. Para mayor comodidad vamos a considerar que las magnitudes $x_1, \ldots, x_n$ están dispuestas en orden creciente de sus valores. Proponemos al lector cerciorarse por sí mismo de que al determinar la distancia por la fórmula (3-3) el valor medio

$$c = \begin{cases} x_{\frac{n+1}{2}} & \text{para } n \text{ impar;} \\ \text{cualquiera de } x_{\frac{n}{2}-1} \text{ a } x_{\frac{n}{2}+1} & \text{para } n \text{ par.} \end{cases} \qquad (3\text{-}25)$$

A veces para ser más concretos para n par se supone que

$$c = \frac{1}{2}\left(x_{\frac{n}{2}-1} + x_{\frac{n}{2}+1}\right). \qquad (3\text{-}26)$$

Si la distancia se determina por la fórmula (3-4), la fórmula para la media es:

$$c = \frac{x_1 + x_n}{2}. \qquad (3\text{-}27)$$

b) Problema de identificación de imágenes

La rama de la cibernética que se ocupa de la identificación de imágenes tiene como objetivo simular una de las propiedades más importantes del cerebro humano, la propiedad de identificar los objetos, fenómenos y situaciones que permiten al hombre orientarse en situaciones ambientales complejas.

Los conceptos fundamentales de la teoría de identificación son los *de clase* e *imagen*. Los diversos objetos o fenómenos del mundo real difieren unos de otros por sus propiedades. Al mismo tiempo los distintos objetos y fenómenos tienen asimismo propiedades generales que permiten agrupar los objetos en ciertos conjuntos o clases. Así pues, los diferentes tipos de automóviles se reúnen en la clase de los automóviles. Las clases pueden ser de diverso grado de generalidad dependiendo de las propiedades que se les atribuyen. Así, la clase de los automóviles es un elemento de la clase más amplia de los medios de transporte. De este modo, se denomina clase un conjunto de objetos o fenómenos reunidos según ciertas propiedades comunes.

En cada situación concreta es necesario operar con un juego finito de clases expresadas por el conjunto finito

$$W = \{A, B, C, \ldots, P\}.$$

Cada clase incluye en sí un conjunto de objetos cuyo número puede ser tan grande como se quiera y cuyas propiedades son sumamente variadas. Sin embargo, en la práctica sólo resulta posible tomar en consideración un número limitado, y con frecuencia muy pequeño, de propiedades diferentes. Llamaremos imagen o modelo del objeto al número total finito de propiedades del objeto, y representaremos la imagen en forma del vector de n dimensiones $x = (x_1, \ldots, x_n)$ cuyas componentes caracterizan cuantitativamente las propiedades de la imagen. Puede servir de ejemplo de imagen la señal de televisión de una fotografía dividida en n celdas independientes. El total del brillo de luminiscencia de las celdas forma un punto en el espacio de n dimensiones constituyendo la imagen de la fotografía.

Puede esperarse (por supuesto, con la condición de que el total de los índices caracterice bastante completamente las propiedades de los objetos), que la totalidad de los puntos correspondientes a las distintas imágenes de una misma clase, va a ocupar en el espacio de imágenes cierto dominio que difiere de los dominios correspondientes a las imágenes de otras clases.

El problema de identificar las imágenes puede considerarse solucionado en principio si en el espacio de imágenes se han hecho pasar planos que lo dividen en dominios cuyos puntos pertenecen a imágenes de clases diferentes. Uno de los métodos posibles de resolución de este problema es el estudio previo de las propiedades de las distintas imágenes de los objetos de cada clase.

Supongamos que existen sólo dos clases de objetos A y B. Designemos por A_0 y B_0 los conjuntos de los modelos de estas clases previamente estudiados. Las representaciones de dichos modelos en el espacio de imágenes pueden tener el aspecto mostrado para el caso de dos dimensiones en la figura 3-4, a, donde los pun-

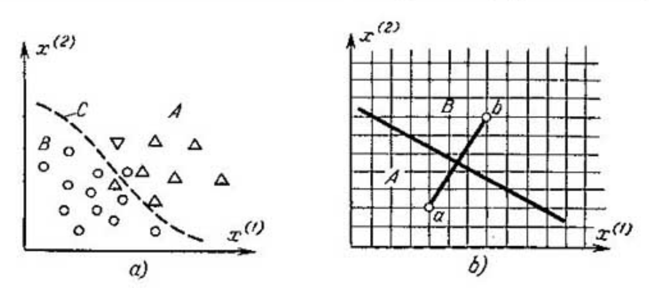


Fig. 3-4. División del espacio de modelos en clases

tos correspondientes a los modelos de A_0 están designados por círculos y de B_0, por triángulos.

Sea x la imagen del nuevo objeto a investigar. Se toman como medida de la cercanía de esta imagen a las clases A y B las magnitudes $S(x, A)$ y $S(x, B)$ que representan, por ejemplo, los cuadrados medios de las distancias del punto x y los puntos correspondientes a las distintas imágenes de las clases A y B. Así pues, si $A_0 = \{a_1, \ldots, a_m\}$ y $B_0 = \{b_1, \ldots, b_l\}$ por la fórmula (3-24) para el valor medio obtenemos:

$$\left.\begin{aligned} S(x, A) &= \frac{1}{m} \sum_{i=1}^{m} d^2(x, a_i); \\ S(x, B) &= \frac{1}{l} \sum_{j=1}^{l} d^2(x, b_j). \end{aligned}\right\} \qquad (3\text{-}28)$$

La superficie que separa las imágenes de ambas clases, en el espacio de imágenes será una superficie C cuyos puntos satisfacen la condición

$$S(x, A) = S(x, B). \qquad (3\text{-}29)$$

Como se ve en la figura 3-4, a, en algunos casos la superficie C será tal que pueda referirse por error la imagen a una clase ajena. Puede servir de criterio de calidad de identificación el número relativo de imágenes identificadas incorrectamente.

Ejemplo 3-5. De cada clase A y B se toma un modelo de $a_1 \in A$ y $b_1 \in B$ que expresan en el espacio bidimensional las propiedades más características de estas clases. Trace una línea que divida el espacio de imágenes en dos dominios correspondientes a las clases A y B.

Se determina la distancia aplicando la fórmula (3-2) por la longitud del segmento que une dos puntos en el espacio de imágenes. Según (3-29), la línea divisoria será aquella todos los puntos de la cual están alejados a una distancia igual de los puntos a_1 y b_1. Esta línea será la perpendicular construida a partir de la mitad del segmento que une los puntos a_1 y b_1 (fig. 3-4, b).

3-5. IMAGENES GEOMETRICAS EN EL ESPACIO MULTIDIMENSIONAL

a) Concepto de hiperesfera

Sea X cierto conjunto de puntos en el espacio multidimensional. Se llama *hiperesfera* la superficie cerrada todos los puntos de la cual están alejados a la misma distancia r de cierto punto fijo a. Por consiguiente, la hiperesfera es el conjunto $X(a, r) \subseteq \mathrm{E}_n$ que se determina como

$$X(a, r) = \{x \in \mathrm{E}_n : d(a, x) = r\}. \qquad (3\text{-}30)$$

El punto fijo a se denomina *centro* y r, *radio* de la hiperesfera. El conjunto

$$B(a, r) = \{x \in \mathrm{E}_n : d(a, x) < r\} \qquad (3\text{-}31)$$

se llama parte interior de la hiperesfera o *esfera abierta*. El conjunto

$$B[a, r] = \{x \in \mathrm{E}_n : d(a, x) \leqslant r\} \qquad (3\text{-}32)$$

se denomina *esfera cerrada*.

La esfera abierta de radio ε con centro en x se llama *vecindad* del punto x y se designa por $O_\varepsilon(x)$.

Ejemplo 3-6. En el espacio $\mathrm{E}_2 = \{x_1, x_2\}$ la ecuación de la hiperesfera se puede escribir en la forma $d^2(a, x) = r^2$. Si la distancia se determina por la fórmula (3-7) la hiperesfera se convierte en esfera: $(x_1 - a_1)^2 + (x_2 - a_2)^2 = r^2$.

b) Conjuntos limitados y finitos

Vamos a denominar conjunto universal el conjunto E_n que representa una reunión infinita de los puntos del espacio de n dimensiones. Designemos por X cierto subconjunto del conjunto E_n. El conjunto X se llama *limitado* si está limitada la distancia entre

dos puntos cualesquiera de dicho conjunto, es decir, si existe un número M tal que para cualesquier $x^{(1)}, x^{(2)} \in X$ se cumple

$$d(x^{(1)}, x^{(2)}) \leqslant M. \qquad (3\text{-}33)$$

Puede demostrarse que la condición necesaria y suficiente para limitar el conjunto X es que el mismo esté en cierta hiperesfera, o sea, la existencia de un punto a y un número r tales que para cualquier $x \in X$ se cumple

$$d(x, a) \leqslant r. \qquad (3\text{-}34)$$

El conjunto X se denomina *finito* si contiene un número finito de puntos. El conjunto finito siempre es limitado.

Ejemplo 3-7. En el espacio E_2 el conjunto $X = \{(x_1, x_2) \in E_2 : x_1^2 + x_2^2 < 1\}$ es un conjunto infinito limitado.

Ejemplo 3-8. En el espacio E_2 el conjunto X que consta de cinco puntos $X = \{(0,0), (0,1), (1,0), (1,1), (1/2, 1/2)\}$, es un conjunto finito.

c) Conjuntos abiertos y cerrados

El punto x se denomina punto *interior* del conjunto X si existe la vecindad $O_\varepsilon(x)$, todos los puntos de la cual pertenecen a X. Se llama conjunto *abierto* al conjunto X todos los puntos del cual son interiores.

Ejemplo 3-9. El intervalo $(a, b) = \{x \in R : a < x < b\}$, en que R es el conjunto de los números reales, constituye un conjunto abierto. Efectivamente, para cualquier $x \in (a, b)$ la vecindad $O_\varepsilon(x)$ en la cual $\varepsilon = \min(x - a, b - x)$ pertenece por completo a (a, b).

Ejemplo 3-10. La esfera abierta $B(a, r)$ es un conjunto abierto. Efectivamente, si $x \in B(a, r)$, entonces $d(a, x) < r$. Supongamos que $\varepsilon = r - d(a, x)$, entonces $O_\varepsilon(x) = B(x, \varepsilon) \subset B(a, r)$.

El punto x se denomina punto *límite* del conjunto X si cualquier vecindad $O_\varepsilon(x)$ de aquél contiene un número infinitamente grande de puntos de X. Si con esto $x \in X$, entonces el mismo se denomina punto límite perteneciente a X. Así pues, todos los puntos interiores del conjunto X son puntos límites pertenecientes a X.

Si x es un punto límite del conjunto X y $x \notin X$, este punto se denomina límite no perteneciente al conjunto X. Tales puntos serán, por ejemplo, los puntos a y b del conjunto (a, b). Un conjunto finito no posee puntos límites.

El conjunto X se llama cerrado si contiene todos sus puntos límites. Pueden servir como ejemplos de conjuntos cerrados cualquier segmento $[a, b]$ de la recta numérica, la esfera cerrada $B[a, r]$ o cualquier conjunto finito.

d) Concepto de hiperplano

El *hiperplano* es la generalización del concepto de plano para el espacio multidimensional. En geometría analítica se demuestra que cualquier ecuación lineal con respecto a las coordenadas de-

fine un plano. Si las coordenadas del punto corriente son x_1, x_2, x_3 la ecuación del plano en el espacio tridimensional es de la forma

$$a_1x_1 + a_2x_2 + a_3x_3 = c. \tag{3-35}$$

Si el juego fijo de números (a_1, a_2, a_3) se entiende como vector a trazado desde el origen de coordenadas a un punto con coordenadas a_1, a_2 y a_3, y como $x = (x_1, x_2, x_3)$ se comprende el vector que determina la posición del punto corriente en el espacio E_3, entonces el primer miembro de la ecuación (3-35) va a constituir el producto escalar de los vectores a y x, y la ecuación

$$ax = c \tag{3-33}$$

será la ecuación vectorial del plano perpendicular al vector $a = (a_1, a_2, a_3)$.

En caso del espacio multidimensional E_n la ecuación (3-63) va a definir el hiperplano perpendicular al vector fijo $a = (a_1, \ldots, a_n)$ cuyas coordenadas corrientes están definidas por el vector $x = (x_1, \ldots, x_n)$. Designando el conjunto de los puntos del hiperplano por $L(x)$ y desarrollando el producto escalar de los vectores a y x llegamos a la siguiente definición de hiperplano;

$$L(x) = \left\{ x \in E_n : \sum_{i=1}^{n} a_i x_i - c = 0 \right\}. \tag{3-37}$$

Los conjuntos

$$L^+ = \left\{ x \in E_n : \sum_{i=1}^{n} a_i x_i - c > 0 \right\} \tag{3-38}$$

y

$$L^- = \left\{ x \in E_n : \sum_{i=1}^{n} a_i x_i - c < 0 \right\} \tag{3-39}$$

se denominan semiespacios abiertos determinados por el hiperplano L.

e) Ecuación del segmento. Concepto de conjunto medio ponderado por elementos

La ecuación de la recta que pasa por los puntos $x^{(1)}$ y $x^{(2)}$ en el espacio E_n se obtiene si se toma en ella el punto x que divide el segmento ($x^{(1)}$, $x^{(2)}$) en la relación w: $1 - w$. Como que la relación de los segmentos de recta será la misma que la relación de sus proyecciones sobre los ejes de coordenadas, para las proyecciones sobre el eje k se llega a la relación

$$\frac{x_k - x_k^{(1)}}{x_k^{(2)} - x_k} = \frac{w}{1 - w}, \tag{3-40}$$

o bien

$$x_k = (1 - w)\, x_k^{(1)} + w x_k^{(2)}. \tag{3-41}$$

Suponiendo que $k = 1, \ldots, n$, obtenemos el sistema de ecuaciones que define la recta:

$$\left. \begin{array}{l} x_1 = (1 - w)\, x_1^{(1)} + w x_1^{(2)}; \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ x_n = (1 - w)\, x_n^{(1)} + w x_n^{(2)}. \end{array} \right\} \qquad (3\text{-}42)$$

Este sistema de ecuaciones puede escribirse de modo más compacto en forma vectorial

$$x = (1 - w)\, x^{(1)} + w x^{(2)}. \qquad (3\text{-}43)$$

La ecuación (3-43) es la de la recta que pasa por los puntos $x^{(1)}$ y $x^{(2)}$, por cierto, como se muestra en la figura 3-5 a diferentes valores de w van a corresponder distintos puntos de dicha recta. En particular, para $0 \leqslant w \leqslant 1$ esta ecuación define el segmento limitado por los puntos $x^{(1)}$ y $x^{(2)}$.

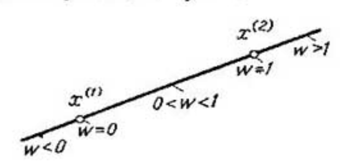


Fig. 3-5. Valores de w en los distintos puntos de la recta

Lo más cómodo es escribir la ecuación del segmento en la forma

$$x = w_1 x^{(1)} + w_2 x^{(2)}; \quad w_1,\ w_2 \geqslant 0; \quad w_1 + w_2 = 1. \qquad (3\text{-}44)$$

La magnitud x definida por (3-44) se denomina *media ponderada* por los elementos $x^{(1)}$ y $x^{(2)}$ con pesos w_1 y w_2. Esta denominación tiene un sencillo sentido físico. Si en los puntos $x^{(1)}$ y $x^{(2)}$ se colocan pesas con los pesos w_1 y w_2, entonces el punto x determina el centro de gravedad del sistema.

El concepto de media ponderada puede extenderse también a un número mayor de puntos. Vamos a llamar media ponderada por los elementos $x^{(1)}, \ldots, x^{(m)}$ a la magnitud x definida por la relación

$$x = \sum_{i=1}^{m} w_i x^{(i)}; \quad w_i > 0, \quad \sum_{i=1}^{m} w_i = 1. \qquad (3\text{-}45)$$

3-6. CONJUNTOS CONVEXOS Y SUS PROPIEDADES

a) Definición de conjunto convexo

Sea X cierto conjunto en el espacio E_n. El conjunto X se denomina *convexo*, si el segmento que une dos puntos cualesquiera de este conjunto pertenece íntegramente a dicho conjunto. En otras

palabras, X es un conjunto convexo si para cualesquier $x^{(1)}, x^{(2)} \in X$ y $w_1, w_2 > 0$, $w_1 + w_2 = 1$ se cumple $w_1 x^{(1)} + w_2 x^{(2)} \in X$.

Los ejemplos de conjuntos convexos en el espacio $E_2 = \{(x, y)\}$ son: $x^2 + y^2 < 1$; $x^2 + y^2 \leqslant 1$; todo el plano E_2; los semiplanos $ax + by - c > 0$ y $ax + by - c < 0$ (fig. 3-6, a). Por otra parte, no son convexos ambos conjuntos definidos por las ecuaciones $x^2 + y^2 \geqslant 1$, $x^2 + y^2 = 1$ (fig. 3-6, b). Efectivamente, el punto (0, 0) que no pertenece a estos conjuntos pertenece al segmento que une los puntos (1, 0) y (—1, 0) de dichos conjuntos.

Teorema 3-1. *La intersección de conjuntos convexos es un conjunto convexo.*

Demostración. Examinemos la intersección $X \cap Y$ de los conjuntos convexos X e Y. Sean x e y dos puntos arbitrarios de $X \cap Y$.

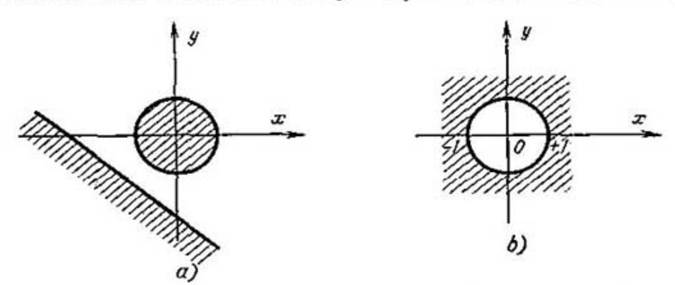


Fig. 3-6. Ejemplos de conjuntos convexos (a) y no convexos (b)

Entonces ellos pertenecen tanto a X como a Y. Por cuanto los conjuntos X e Y son convexos, el segmento que une los puntos x e y pertenece íntegramente tanto a X como a Y, y por eso, a la intersección $X \cap Y$. Por consiguiente, $X \cap Y$ es un conjunto convexo.

b) Envoltura convexa de un conjunto finito

Sea $A = \{a_1, \ldots, a_m\}$ el conjunto finito de puntos en el espacio E_n. El conjunto finito no es convexo. No obstante, los puntos del conjunto finito A pueden ser elementos de algunos conjuntos convexos, por ejemplo S_1, S_2, etc., como se muestra en la figura 3-7. En este caso A es un subconjunto de los conjuntos convexos S_1, $S_2 \ldots$

Se denomina envoltura convexa co(A) del conjunto A la intersección de todos los conjuntos convexos cuyos subconjuntos es A. En particular, si $A \subset S_1$ y $A \subset S_2$, entonces co$(A) \subseteq S_1 \cap S_2$.

De la definición dada se infiere que la envoltura convexa co(A) es el conjunto convexo menor que contiene A. En efecto, si hubiera un conjunto convexo S tal que $A \subset S$ y $S \subseteq$ co(A), entonces se

cumpliría que $\text{co}(A) \subseteq S \cap \text{co}(A)$, es decir, $\text{co}(A) \subseteq S$. De aquí se deduce que $S = \text{co}(A)$.

Teorema 3-2. *La envoltura convexa del conjunto finito A es el conjunto de las medias ponderadas por elementos del conjunto A.*

Sean $A = \{a_1, \ldots, a_m\}$ un conjunto finito y w_i, $i = 1, \ldots, m$ números reales tales que

$$w_i \geqslant 0; \quad \sum_{i=1}^{m} w_i = 1. \tag{3-46}$$

La magnitud

$$x = \sum_{i=1}^{m} w_i a_i \tag{3-47}$$

se denomina media ponderada por elementos del conjunto A. El conjunto S cuyos elementos son las magnitudes x determinadas por la fórmula (3-47) para todas las w_i posibles en el marco de las limitaciones (3-46), constituye el conjunto de las medias ponderadas por elementos del conjunto A. Se requiere demostrar que

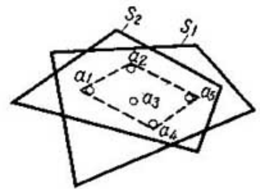


Fig. 3-7. Para la obtención de la envoltura convexa de un conjunto finito

$$S = \text{co}(A). \tag{3-48}$$

Demostración. Para que el conjunto S sea la envoltura convexa del A, aquél deberá ser, en primer lugar, convexo y, en segundo, el menor conjunto convexo que contiene el conjunto A. Demostremos que se cumplen ambas condiciones.

1. Tomemos dos sistemas arbitrarios de pesos w_i y w'_i, que determinan dos puntos del conjunto S:

$$x = \sum_{i=1}^{m} w_i a_i, \quad x' = \sum_{i=1}^{m} w'_i a_i. \tag{3-49}$$

Tomemos el punto arbitrario x'' del segmento que une x y x':

$$x'' = (1 - w)\, x + wx' = \sum_{i=1}^{m} [(1 - w)\, w_i + ww'_i]\, a_i = \sum_{i=1}^{m} w''_i a_i, \tag{3-50}$$

donde

$$w''_i = (1 - w)\, w_i + ww'_i. \tag{3-51}$$

La magnitud w''_i no es negativa ya que representa la media ponderada de dos números no negativos w_i y w'_i. Además,

$$\sum_{i=1}^{m} w''_i = (1 - w) \sum_{i=1}^{m} w_i + w \sum_{i=1}^{m} w'_i = 1. \tag{3-52}$$

De este modo x'' es la media ponderada por elementos del conjunto A, es decir, $x'' \in S$. Por consiguiente, S es un conjunto convexo.

2. Sea T un conjunto convexo arbitrario que contiene A. Demostremos que el mismo también contiene S.

Por hipótesis todos los elementos a_i están contenidos en T. Más adelante, el elemento

$$x_2 = \frac{w_1}{w_1 + w_2} a_1 + \frac{w_2}{w_1 + w_2} a_2 \tag{3-53}$$

es la media ponderada de a_1, $a_2 \in T$, de modo que $x_2 \in T$. Procediendo exactamente igual obtenemos que

$$x_3 = \frac{w_1 + w_2}{w_1 + w_2 + w_3} x_2 + \frac{w_3}{w_1 + w_2 + w_3} a_3 = \\ = \frac{w_1 a_1}{w_1 + w_2 + w_3} + \frac{w_2 a_2}{w_1 + w_2 + w_3} + \frac{w_3 a_3}{w_1 + w_2 + w_3} \tag{3-54}$$

es la media ponderada de x_2, $a_3 \in T$, de manera que $x_3 \in T$. Prosiguiendo el examen de esta secuencia de puntos, hallamos que el punto

$$x_m = \frac{w_1 + \dots + w_{m-1}}{w_1 + \dots + w_m} x_{m-1} + \frac{w_m}{w_1 + \dots + w_m} a_m = \sum_{i=1}^{m} w_i a_i \in S \tag{3-55}$$

pertenece asimismo al conjunto T. De esta suerte, cualquier $x \in S$ también pertenece a T, o sea, $S \subseteq T$. Pero T es un conjunto convexo arbitrario que contiene A. Consecuentemente, S es el conjunto convexo menor que contiene A.

El teorema dado permite determinar qué aspecto tiene la envoltura convexa del conjunto finito. Consideremos el caso bidimensional. En la figura 3-8 se muestran el conjunto finito $A = \{a_1, a_2, a_3, a_4, a_5\}$ y el polígono convexo cuyos vértices son los elementos del conjunto A. Es fácil ver que cualquier punto en un vértice, un lado o en el interior de un polígono convexo puede representarse como la media ponderada de, por lo menos, tres de sus vértices.

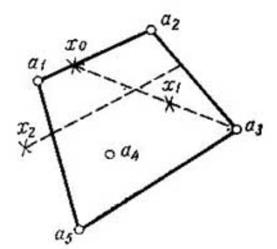


Fig. 3-8. Envoltura de un conjunto finito

Así pues, $x_0 = \text{med. pon}(a_1, a_2)$, $x_1 = \text{med. pon}(x_0, a_3) = \text{med. pon}(a_1, a_2, a_3)$. Al mismo tiempo el punto x_2 que está fuera del polígono no puede ser la media ponderada de dos puntos interiores ningunos y esto significa que no puede expresarse como la media ponderada de los vértices del polígono. De este modo, la envoltura convexa del conjunto finito A en el plano es el polígono cuyos vértices son los elementos del conjunto A.

Esta deducción se generaliza fácilmente para el caso de un espacio arbitrario E_n. La envoltura convexa del conjunto finito A en el espacio E_n es el polígono convexo cuoys vértices son los elementos del conjunto A. Cualquier punto interior o límite de tal polígono puede ser representado como la media ponderada de no más de $(n - 1)$ vértices.

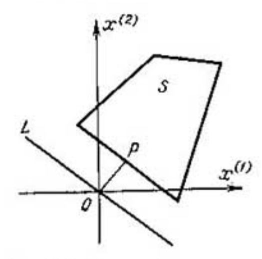


Fig. 3-9. Para la construcción del hiperplano de apoyo

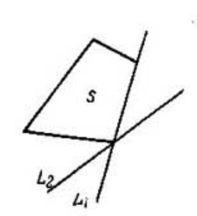


Fig. 3-10. Rectas de apoyo al conjunto convexo

Para concluir enunciemos sin demostración dos teoremas que serán utilizados para fundamentar una serie de preceptos de la programación lineal y de la teoría de los juegos.

Teorema 3-3 (teorema del hiperplano de apoyo). *Sea S un conjunto convexo y O un punto arbitrario que no pertenece a S. El teorema afirma que existe un hiperplano L que pasa por O, tal que S está en uno solo de los semiplanos determinados por L* (para los espacios bidimensional y tridimensional esto resulta intuitivamente claro por la fig. 3-9).

Vamos a desplazar el hiperplano en el espacio acercándolo a S hasta que L y S vayan a tener, por lo menos, un punto común quedando S, no obstante, en uno de los semiplanos determinados por el hiperplano L. Este se denomina hiperplano de apoyo para el conjunto convexo S.

Fig. 3-11. Hiperplano divisor

Si el conjunto convexo S es la envoltura del conjunto finito A, entonces aquél representa un polígono convexo. En este caso el plano de apoyo va a pasar por el vértice, arista o borde del polígono (fig. 3-10). En cualquiera de dichas variantes resulta que al menos uno de los vértices del polígono está en el hiperplano de apoyo.

Teorema 3-4 (teorema del hiperplano divisor). *Si S y T son conjuntos convexos no intersecados y, por lo menos, uno de ellos está limitado, existe un hiperplano tal que los conjuntos S y T*

están en los diferentes semiespacios determinados por dicho hiperplano. En otras palabras, existe el hiperplano $px = c$, donde $p = (p_1, \ldots, p_n)$ es cierto conjunto ordenado de números reales, y $x = (x_1, \ldots, x_n)$ es un punto arbitrario, tal que $px > c$, si $x \in S$ y $px < c$, si $x \in T$. La ilustración geométrica de este teorema se muestra en la figura 3-11.

PROBLEMAS PARA EL CAPITULO 3

3-1. En el espacio tridimensional están dados los puntos $x = (1, 0, 2)$, $y = (3, 4, 0)$, $z = (1, 2, 3)$. Comprobar para estos puntos el axioma del triángulo en los espacios $C_2^{(3)}$, $C_1^{(3)}$, $C^{(3)}$.

3-2. Dibujar en el plano (x, y) la circunferencia con centro en a y radio r en los espacios $C_2^{(3)}$, $C_1^{(3)}$, $C^{(3)}$.

3-3. Demostrar las fórmulas (3-25) y (3-27) para determinar la media.

3-4. Determinar el cambio cualitativo de la línea que separa las clases A y B en el ejemplo 3-5, si en la clase A hay además un segundo modelo a_2. Analizar el caso en que $a_2 = (2, 1)$, $a_2 = (2, 3)$ y $a_2(6, 3)$.

3-5. ¿Qué aspecto va a tener la línea que separa las clases A y B en la figura 3-4, *b* si la distancia se determina:

a) por la fórmula (3-3),

b) por la fórmula (3-4).

3-6. Demostrar que la esfera abierta $B(0, r)$ en el espacio E_n es un conjunto convexo.

3-7. ¿Cuáles de los conjuntos $[0, 3]$, $[5, 7]$, $[0, 3] \cup [5, 7]$ en el espacio de números reales R son convexos y cuáles no lo son?

3-8. Demostrar que el punto x_1 en la figura 3-8 es la media ponderada de los vértices a_1, a_2 y a_3.

Capítulo cuarto
ELEMENTOS DE ALGEBRA LOGICA

4-1. OPERACIONES LOGICAS

a) Concepto de proposición

Como vimos antes pueden presentarse los conjuntos por dos métodos: de enumeración y de descripción, o sea, indicando la propiedad que poseen los elementos del conjunto. El método descriptivo para presentar los conjuntos vincula la teoría de conjuntos con la teoría de las proposiciones que constituye la primera y más sencilla parte de la disciplina científica llamada lógica matemática.

Llamaremos proposición a cada afirmación que puede ser verdadera o falsa. Vamos a examinar las proposiciones con respecto a los elementos de cierto conjunto universal I. Los diversos elementos de este conjunto van a poseer diferentes propiedades y en concordancia con esto pueden formar distintos grupos que constituyen subconjuntos del conjunto I. Así pues, si I es el conjunto de estudiantes en un grupo, sus subconjuntos pueden ser: X, el conjunto de estudiantes sobresalientes; Y, el conjunto de estudiantes que viven en la residencia estudiantil; Z, el conjunto de estudiantes becados, etc.

Después que se hayan distinguido las propiedades que poseen los diversos subconjuntos pueden hacerse determinadas afirmaciones respecto a si uno u otro elemento de I tiene las propiedades requeridas. Estas afirmaciones serán proposiciones: "es estudiante sobresaliente", "vive en la residencia de estudiantes" y "es becario".

En lo sucesivo nos abstraeremos de todas las propiedades de las proposiciones a excepción de una: cada una de las proposiciones puede ser verdadera o falsa con respecto al elemento considerado del conjunto universal I. Así bien, la proposición "el estudiante Ivanov es sobresaliente" es cierta si el estudiante Ivanov pertenece al subconjunto X, y falsa si no pertenece a éste.

Por eso el subconjunto X se denomina *conjunto de veracidad* para la proposición "es estudiante sobresaliente". Los conjuntos de veracidad para las proposiciones "vive en la residencia de estudiantes" y "es becario" serán, respectivamente, los conjuntos Y y Z.

El conjunto de veracidad de cierta proposición puede resultar vacío. En este caso la proposición se denomina idénticamente falsa. Así, para el grupo estudiantil será idénticamente falsa la proposición "es mayor de cincuenta años". Puede suceder que el conjunto de veracidad de cierta proposición coincida con el conjunto universal I. En este caso la proposición se llama idénticamente verdadera. Para el grupo estudiantil será idénticamente verdadera la proposición "es menor de cincuenta años".

b) Proposiciones simples y compuestas

Vamos a designar las proposiciones con letras minúsculas del alfabeto romano y atribuir a cada una de ellas valores numéricos de 1 y 0 según sea verdadera o falsa la proposición. Supongamos por ejemplo que x denota la proposición "es estudiante sobresaliente". Sus valores numéricos son:

$$x = \begin{cases} 1, \text{ si } x \text{ es verdadera, o sea, } x \in X; \\ 0, \text{ si es falsa, o sea, } x \notin X. \end{cases} \tag{4-1}$$

Las proposiciones x, y, z a las cuales corresponden los conjuntos de veracidad sencillos X, Y, Z se denominan proposiciones *sencillas*. No obstante, puede resultar un conjunto de veracidad el conjunto Q obtenido de los conjuntos X, Y, Z por medio de cualquier operación algebraica con dichos conjuntos. En tal caso va a corresponder al conjunto de veracidad Q la proposición q llamada proposición *compuesta*. Por ejemplo, al conjunto de veracidad $Q = X \cap Y$ que posee tanto la propiedad X (estudiante sobresaliente) como la propiedad Y (vive en la residencia de estudiantes), va a corresponder la proposición compuesta "es estudiante sobresaliente y vive en la residencia estudiantil".

En el ejemplo dado obtuvimos una proposición compuesta enlazando dos proposiciones mediante la conjunción "y". Sería posible obtener una proposición compuesta utilizando otras cópulas: "o", "si, entonces", etc. De una proposición puede obtenerse una nueva negándola. A cada una de estas nuevas proposiciones van a corresponder sus conjuntos de veracidad en el conjunto universal I, es decir, las proposiciones compuestas pueden ser también verdaderas o falsas, o sea, tomar los valores numéricos 1 y 0.

Considerando las proposiciones como magnitudes que toman valores 1 y 0 pueden determinarse con las mismas las operaciones que permiten obtener de las proposiciones dadas otras nuevas. Estas operaciones van a expresar las cópulas citadas más arriba que se emplean en el lenguaje corriente.

Las operaciones realizadas con las proposiciones se denominan *operaciones lógicas*. El total de las operaciones lógicas que se examinan más adelante recibió el nombre de *álgebra de proposiciones* o *álgebra booleana* (de Boole).

c) Presentación de las operaciones lógicas

Supongamos que tenemos varias proposiciones sencillas $x_1, \ldots, x_N$ cada una de las cuales puede ser verdadera o falsa, o sea, adquirir valores numéricos 1 y 0. Podemos considerar el total de estas proposiciones como el cortejo $(x_1, \ldots, x_N)$.

Tabla 4-1

Funciones de dos variables lógicas

x	y	0	x	y	xy	$x+y$	$x\oplus y$	$x\to y$	$y\to x$	1	$\bar{x}$	$\bar{y}$	$\overline{xy}$	$\overline{x+y}$	$\overline{x\oplus y}$	$\overline{x\to y}$	$\overline{y\to x}$
0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0
0	1	0	0	1	0	1	1	1	0	1	1	0	1	0	0	0	1
1	0	0	1	0	0	1	1	0	1	1	0	1	1	0	0	1	0
1	1	0	1	1	1	1	0	1	1	1	0	0	0	0	1	0	0

Supongamos que con estas proposiciones hemos realizado cierta operación lógica como resultado de la cual obtuvimos una nueva proposición q que puede ser asimismo verdadera o falsa, es decir, puede tomar los valores 1 y 0. En este caso a cada combinación de valores $x_1, \ldots, x_N$ va a corresponder un valor determinado $q \in \{1, 0\}$. En consecuencia, la operación lógica puede considerarse como el reflejo f del conjunto de los valores del cortejo $(x_1, \ldots, x_N)$ en el conjunto de los valores q

$$f: (x_1, \ldots, x_N) \to \{1, 0\}.$$

Tabla 4-2

$q = \bar{x}y + z$

x	y	z	q
0	0	0	0
1	0	0	0
0	1	0	1
0	0	1	1
1	1	0	0
1	0	1	1
0	1	1	1
1	1	1	1

Si este reflejo es unívoco el mismo define la función

$$q = f(x_1, \ldots, x_N), \qquad (4\text{-}2)$$

que se denomina *función booleana*. Como se infiere de lo dicho, en las funciones booleanas tanto los argumentos, como las funciones propiamente dichas sólo pueden tomar dos valores diferentes que designamos por 1 y 0.

Se emplean tres métodos para representar las funciones booleanas:

1. La fórmula, que indica en forma explícita la secuencia de operaciones lógicas que se realizan con las proposiciones $x_1, \ldots, x_N$ y que tiene la forma de relación (4-2).

2. La tabla, que señala los valores de veracidad de la proposición compuesta q según los valores de veracidad de las proposi-

ciones iniciales. En la parte izquierda de la tabla se enumeran todas las combinaciones posibles de valores de veracidad $x_1, \ldots, x_N$, y en la parte derecha, los valores de veracidad de la proposición compuesta q. Si hay N proposiciones iniciales, el número de líneas de la tabla será de 2^N. Pueden servir como ejemplos de ellas las tablas 4-1 y 4-2.

3. El esquema lógico que representa la designación gráfica de referencia convencional de la operación lógica que tiene el aspecto mostrado en la figura 4-1.

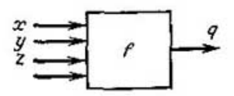


Fig. 4-1. Designación convencional de la operación lógica

En automática y técnica de cálculo se acostumbra a presentar las diferentes proposiciones en forma de señales que tienen dos niveles, o en forma de dispositivos capaces de ocupar dos posiciones (relé, trigger, tubo electrónico, transistor, etc.). A estos dos niveles de las señales o dos estados del dispositivo se atribuyen los valores numéricos 1 y 0 que determinan los valores de veracidad de las proposiciones correspondientes. Con esta concepción el esquema lógico va a constituir un convertidor de señales que puede utilizarse con fines de mando de diversos procesos.

Ejemplo 4-1 Para mantener la temperatura Θ de un termostato en el nivel deseado Θ_0 se emplea el elemento calentador *EC* (fig. 4-2) que puede funcionar en dos regímenes: "conectado" y "desconectado". El mando de la conexión del elemento calentador se realiza por la regla:

conectar el *EC* si $\Theta < \Theta_0$;

desconectar el *EC* si $\Theta > \Theta_0$.

Fig. 4-2. Regulación de la temperatura en un termostato

Tomemos la proposición "temperatura mayor de Θ_0" como proposición inicial x y la proposición "conectar el *EC*", como proposición compuesta q. Entonces la estructura lógica del circuito de mando del elemento calentador va a tener la forma:

$$q = \begin{cases} 1, & \text{si } x = 0; \\ 0, & \text{si } x = 1. \end{cases}$$

En la figura 4-2 se muestra la realización física de la ley de mando descrita. La señal x se expresa por la altura de la columna de mercurio del termómetro T que con $\Theta \geqslant \Theta_0$ cierra el contacto del relé R, conectando así el elemento calentador.

4-2. ALGEBRA DE PROPOSICIONES

a) Negación

Sea x la proposición cuyo conjunto de veracidad es X. Designemos por $\bar{x}$ (se lee "no es x") una nueva proposición que tiene el conjunto de veracidad $\bar{X}$ y que se denomina *negación* o *inversión* de X.

La relación entre x y $\bar{x}$ puede determinarse utilizando el diagrama de Euler — Venn dado en la figura 4-3, *a*. Según la definición de negación, $\bar{x} = 1$ en la zona $\bar{X}$, donde $x = 0$ y $\bar{x} = 0$ en la zona X donde $x = 1$. Esto puede escribirse en forma de las siguientes reglas de inversión:

$$\bar{1} = 0, \ \bar{0} = 1. \qquad (4\text{-}3)$$

En la figura 4-3, *b* se muestra la representación de la operación de negación en forma del esquema lógico denominado *inversor*.

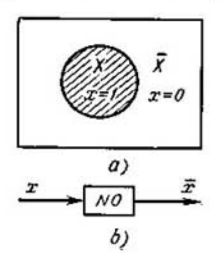


Fig. 4-3. Diagrama de Euler — Venn y designación convencional de la operación de negación

b) Adición lógica

Sean x e y proposiciones que tienen conjuntos de veracidad X e Y, respectivamente. Designemos por $x + y$ (a veces se escribe $x \vee y$ y se lee x o y) la nueva proposición que tiene el conjunto de veracidad $X \cup Y$ y que se denomina *suma lógica* o *disyunción* de x e y.

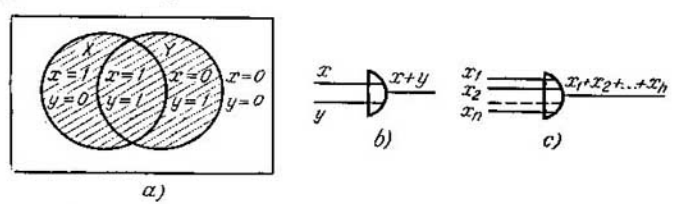


Fig. 4-4. Diagrama de Euler — Venn y designación convencional de la operación de adición lógica

Las relaciones entre los valores numéricos x, y y $x + y$ se toman del diagrama de Venn en la figura 4-4, *a*. El conjunto de veracidad de la proposición $x + y$ es la zona rayada en correspondencia con la cual

$$x + y = \begin{cases} 0, \text{ si } x = 0 \text{ e } y = 0; \\ 1, \text{ en todos los demás casos,} \end{cases} \qquad (4\text{-}4)$$

Es cómodo escribir las relaciones (4-4) en forma de reglas que recuerdan algo las reglas de la suma aritmética habitual:

$$\begin{aligned} &0+0=0; \quad 1+0=1; \\ &0+1=1; \quad 1+1=1. \end{aligned} \tag{4-5}$$

En la figura 4-4, *b* se muestra el esquema lógico utilizado para representar la operación de adición lógica que se denomina *esquema de reunión.*

La regla de adición lógica se extiende fácilmente al caso de tres o más proposiciones. En el caso general la suma lógica de n proposiciones se determina por la regla

$$x_1+x_2+\ \dots\ +x_n=\begin{cases} 0, \text{ si } x_1=x_2=\ \dots\ x_n=0; \\ 1, \text{ en todos los demás casos.} \end{cases} \tag{4-6}$$

El esquema lógico para este caso se muestra en la figura 4-4, *c*.

c) Multiplicación lógica

Sean x e y las proposiciones que tienen los conjuntos de veracidad X e Y, respectivamente. Designemos por xy (a veces se escribe $x \wedge y$ y se lee x e y) la nueva proposición que tiene el conjunto de

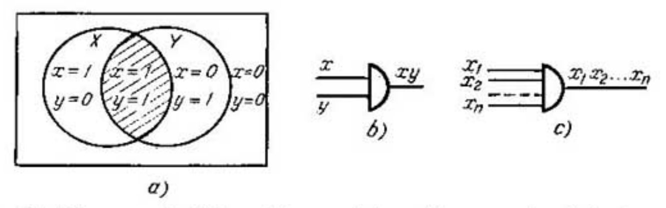


Fig. 4-5. Diagrama de Euler — Venn y designación convencional de la operación de multiplicación lógica

veracidad $X \cap Y$ y se denomina *producto lógico* o *conjunción* de x e y.

El conjunto de veracidad de las proposiciones xy está dado por la zona rayada del diagrama de Euler — Venn en la figura 4-5, *a* al examinar la cual obtenemos:

$$xy=\begin{cases} 1, \text{ si } x=1 \text{ e } y=1; \\ 0, \text{ en todos los demás casos.} \end{cases} \tag{4-7}$$

Esta relación puede representarse en forma de reglas de multiplicación lógica que coinciden con las reglas de multiplicación aritmética:

$$0\cdot 0=0, \quad 0\cdot 1=0, \quad 1\cdot 0=0, \quad 1\cdot 1=1. \tag{4-8}$$

En la figura 4-5, *b* se muestra el esquema lógico empleado para representar la operación de multiplicación lógica y denominado *esquema de coincidencia.*

La regla de multiplicación lógica se extiende al caso de tres o más proposiciones. Así pues, el producto lógico de *n* proposiciones se determina por la regla

$$x_1 x_2 \ldots x_n = \begin{cases} 1, & \text{si } x_1 = x_2 = \ldots = x_n = 1; \\ 0, & \text{en todos los demás casos.} \end{cases} \tag{4-9}$$

El esquema lógico para este caso se presenta en la figura 4-5, *c*.

Una de las variantes del esquema de multiplicación lógica es la operación del producto lógico de una proposición x por la inversión de otra $\bar{y}$, la que puede representarse en la forma:

$$x\bar{y} = \begin{cases} x, & \text{si } y = 0 \\ 0, & \text{si } y = 1. \end{cases} \tag{4-10}$$

Como dicha operación se encuentra con bastante frecuencia, para representarla es conveniente emplear el esquema lógico especial dado en la figura 4-6. Este esquema lógico tiene el aspecto

Fig. 4-6. Esquema de inhibición

del esquema de coincidencia con la diferencia de que la entrada de inversión y no termina una flecha sino en un circulito. Como se ve en (4-10), la aplicación de una señal $y = 1$ a dicha entrada resulta que ella como si cerrara el camino para el paso de la señal x a la salida del circuito. Por eso la entrada y se llama *entrada de inhibición* y el esquema de la figura 4-6 recibió el nombre de *esquema de inhibición.* La operación lógica propiamente dicha correspondiente a las correlaciones (4-10) se denomina *operación de inhibición.*

d) Funciones booleanas

Algunas proposiciones a las cuales se atribuyen valores numéricos de veracidad de 1 y 0 pueden considerarse como unas variables binarias. En lo adelante vamos a denominarlas variables lógicas. Las operaciones de negación, adición y multiplicación lógicas constituyen funciones de las variables lógicas; con esto, la operación de negación es función de una variable lógica y las opera-

ciones de adición y multiplicación lógicas son funciones de dos o más variables lógicas.

Como se ha indicado ya, la particularidad de estas funciones consiste en que ellas sólo pueden tomar dos valores 0 y 1.

Las funciones en las cuales tanto los argumentos como las mismas funciones sólo pueden tomar dos valores distintos se denominan funciones booleanas. Naturalmente, además de las funciones consideradas pueden haber otras funciones booleanas. Sin embargo, puede demostrarse que la cantidad de diversas funciones booleanas de un número finito de variables lógicas, es finita.

En efecto, supongamos que queremos formar todas las funciones booleanas posibles de m variables lógicas $x, y, z \ldots$ Llamaremos *juego* al total de los valores numéricos de estas variables. Es posible obtener por todo 2^m juegos.

Sea $f(x, y, z \ldots)$ una función booleana. A cada uno de los juegos corresponde un valor numérico determinado (1 ó 0) de la función booleana, así que ésta puede considerarse como un juego de 2^m variables. Las distintas funciones booleanas van a diferir por los valores de las variables lógicas en este juego. Por cuanto cada variable puede tomar solamente dos valores, el número de diversas funciones booleanas de m argumentos es:

$$B(m) = 2^{2^m}. \qquad (4\text{-}11)$$

Con $m = 1$ podemos obtener $2^2 = 4$ funciones booleanas distintas, a saber: las constantes 1 y 0, la propia función x y su negación $\bar{x}$.

Con $m = 2$ hay $2^4 = 16$ funciones booleanas diferentes. En la tabla 4-1 se dan los valores de estas funciones y sus designaciones convencionales dispuestas de tal modo que la segunda mitad de la tabla se obtiene a partir de la primera utilizando la operación de negación. Además de las funciones examinadas anteriormente, forman parte de la tabla 4-1 las siguientes funciones:

$x \oplus y$, adición por el módulo 2;
$x \rightarrow y$, $y \rightarrow x$, implicación u operación de secuencia lógica;
$\overline{xy} = x/y$, función de Sheffer;
$\overline{x + y}$, función de Peirce (función de Webb);
$\overline{x \oplus y} = x \sim y$, equivalencia lógica;
$\overline{x \rightarrow y}$, $\overline{y \rightarrow x}$, función de inhibición.

Es importante destacar que todas las funciones booleanas de dos variables lógicas examinadas pueden representarse con ayuda de tres operaciones lógicas consideradas anteriormente: negación, multiplicación lógica y adición lógica. Así pues, mediante la comprobación directa en todos los juegos (x, y) es fácil cerciorarse de

que se cumplen las siguientes relaciones:

$$\left.\begin{array}{ll}(a) & x\oplus y=\bar{x}y+\bar{y}x;\\ (b) & x\rightarrow y=\bar{x}+y;\\ (c) & y\rightarrow x=x+\bar{y};\\ (d) & \overline{x\oplus y}=x\sim y=xy+\bar{x}\bar{y};\\ (e) & \overline{x\rightarrow y}=x\bar{y};\\ (f) & \overline{y\rightarrow x}=y\bar{x}.\end{array}\right\} \qquad (4.12)$$

El juego de operaciones lógicas se llama juego completo, si el mismo permite representar cualquier función de Boole. Además del juego ya examinado que consta de las operaciones de negación, multiplicación lógica y adición lógica pueden haber también otros juegos completos. Así, puede servir de juego completo una sola operación, la función de Sheffer $\overline{xy}$ designada con frecuencia por x/y, de lo que es fácil cerciorarse representando en forma de esta operación las tres operaciones del juego completo arriba mencionadas:

$$\bar{x}=\overline{xx}=x/x;$$

$$xy=\overline{\overline{xy}}=\overline{x/y}=(x/y)/(x/y);$$

$$x+y+\overline{\bar{x}\bar{y}}=\bar{x}/\bar{y}=(x/x)/(y/y).$$

También es fácil mostrar que la función de Peirce constituye asimismo un juego completo.

e) Leyes e identidades del álgebra de proposiciones

Vamos a llamar fórmulas booleanas (de Boole) y designar $A(x, y, z\ldots)$, $B(x, y, z)$ las expresiones formadas por un número finito de variables lógicas $x, y, z\ldots$, por signos de las operaciones lógicas de negación, multiplicación lógica y adición lógica y también por las constantes 0 y 1. Para la comprensión unívoca de las fórmulas booleanas cada nueva operación lógica deberá ir acompañada por la introducción de un par de paréntesis. Sin embargo, para evitar que las fórmulas sean excesivamente complicadas y disminuir el número de paréntesis empleados se introduce por analogía con el álgebra elemental el concepto de prioridad de las operaciones: primeramente deben realizarse las operaciones de negación, después, de multiplicación lógica y, por último, las operaciones de adición lógica.

Cada fórmula booleana puede considerarse como la representación de cierta función booleana de las variables $x, y, z\ldots$, cuyo valor en un juego concreto de variables es fácil de obtener si se colocan los valores de las variables en este juego (0 ó 1) en la fórmula booleana y se realizan las operaciones lógicas señaladas.

Para unas mismas variables lógicas pueden obtenerse distintas fórmulas lógicas entre las cuales pueden encontrarse fórmulas tales como, por ejemplo, A y B que dan los mismos valores de las funciones booleanas en todos los juegos iguales de las variables lógicas $x, y, z \ldots$, es decir, satisfacen la condición

$$A(x, y, z \ldots) = B(x, y, z \ldots). \tag{4-13}$$

Una de las tareas del álgebra booleana es establecer correlaciones idénticas del tipo (4-13).

Para comprobar si son correctas las relaciones idénticas del álgebra booleana pueden calcularse sencillamente los valores de las funciones en los primero y segundo miembros de la identidad en todos los 2^m juegos de variables.

Otro método para establecer la identidad de dos fórmulas booleanas es determinar sus conjuntos de veracidad. Si para dos fórmulas booleanas los conjuntos de veracidad coinciden, estas fórmulas dan una misma función booleana.

En el capítulo 1 fue demostrada una serie de identidades del algebra de los conjuntos. Cada una de esta identidades determina cierta identidad para las fórmulas booleanas. Examinemos, por ejemplo, la siguiente identidad del álgebra de los conjuntos:

$$(X \cap Y) \cup Z = (X \cup Z) \cap (Y \cup Z).$$

El primer miembro de esta identidad es el conjunto de veracidad para la fórmula booleana $xy + z$, y el segundo miembro, para $(x + z)(y + z)$. Por cuanto los conjuntos de veracidad coinciden para ambas fórmulas booleanas, estas fórmulas dan una misma función booleana, es decir, se cumple la identidad

$$xy + z = (x + z)(y + z).$$

Esta identidad es notable por el hecho de que no tiene analogía en el álgebra ordinaria sobre lo que se hizo énfasis en el capítulo 1.

Proponemos al lector cerciorarse por sí mismo de que se cumplen las leyes e identidades de la lógica matemática que siguen más abajo, confrontándolas con las identidades correspondientes del álgebra de los conjuntos o comprobando los valores de veracidad en los primero y segundo miembros en todos los juegos de variables lógicas.

Leyes del álgebra booleana

a) $x + y = y + x$;
b) $xy = yx$; } ley conmutativa;

c) $(x + y) + z = x + (y + z)$;
d) $(xy)z = x(yz)$; } ley asociativa;

e) $(x + y)z = xz + yz$;
f) $xy + z = (x + z)(y + z)$. } ley distributiva.

Identidades del álgebra booleana

a) $x\bar{x} = 0$; g) $xx = x$;

b) $x + \bar{x} = 1$; h) $x + x = x$;

c) $x \cdot 1 = x$; i) $\bar{\bar{x}} = x$;

d) $x + 1 = 1$; j) $\overline{x + y} = \bar{x}\bar{y}$;

e) $x \cdot 0 = 0$; k) $\overline{xy} = \bar{x} + \bar{y}$.

f) $x + 0 = x$;

4-3. SINTESIS DE LOS ESQUEMAS COMBINATORIOS

a) Concepto de esquema combinatorio

Vamos a entender por *esquema combinatorio* un dispositivo técnico que sirve para convertir la información discreta (discontinua) y tiene n circuitos de entrada y m circuitos de salida. Las señales suministradas a los circuitos de entrada y tomadas de los circuitos de salida sólo pueden tener los valores 1 y 0. Estas señales pueden ser, por ejemplo, los niveles alto y bajo de la tensión.

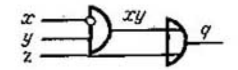


Fig. 4-7. Ejemplo de esquema lógico combinatorio

De esta manera, el esquema combinatorio convierte cierta palabra de entrada de n letras del alfabeto binario en una palabra de salida de m letras del mismo alfabeto. Al hacerlo se supone que la palabra de salida se recibe en la salida en el mismo momento en que se suministra a la entrada la palabra de entrada, es decir, en el esquema combinatorio no hay retardo de señales.

Frecuentemente se utilizan en calidad de elementos, con los cuales se construye el esquema combinatorio, esquemas lógicos elementales para las operaciones de negación, multiplicación lógica y adición lógica. En tal caso puede obtenerse el esquema combinatorio conectando en serie los circuitos de salida de unos esquemas lógicos elementales a los circuitos de entrada de otros esquemas lógicos elementales y evitando conectar a un mismo circuito de entrada varios circuitos de salida. Si el esquema combinatorio tiene un conjunto finito de señales de entrada $x, y, z \ldots$ y una salida única q, entonces el proceso descrito de conexión en serie de los esquemas lógicos elementales corresponde al proceso de formación de la fórmula lógica $f(x, y, z \ldots)$ con ayuda de las operaciones realizadas con los esquemas lógicos elementales que se utili-

zan. Así pues, a la fórmula lógica del tipo

$$q = \bar{x}y + z \tag{4-14}$$

va a corresponder el esquema combinatorio dado en la figura 4-7 y en la tabla 4-2. En diversos casos es conveniente armar un es-

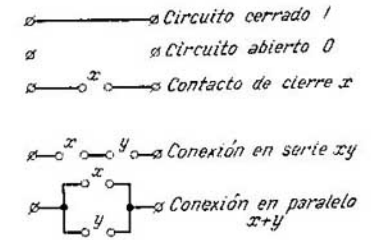


Fig. 4-8. Representación de las operaciones lógicas elementales con ayuda de contactos

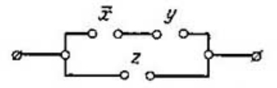


Fig. 4-9. Esquema de conmutación para la operación $\bar{x}y + z$

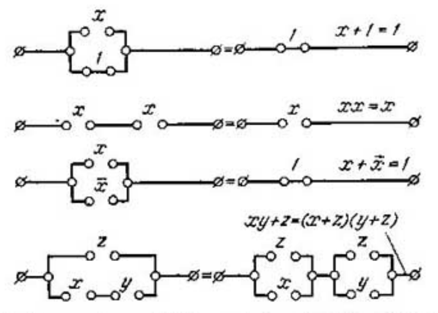


Fig. 4-10. Conversiones idénticas de los circuitos de conmutación

quema combinatorio con los elementos que realizan la función de Sheffer o la de Peirce.

Sirven como uno de los métodos de realización material de los esquemas combinatorios los esquemas de relé de contactos que permiten efectuar las operaciones lógicas mediante el cierre y la apertura de los contactos de varios circuitos. Designemos por 1 el circuito cerrado y por 0 el circuito abierto. Para gobernar el estado

del circuito introducimos en él los contactos cuyo estado se determina por el valor de las variables de entrada x, y, z ... Es fácil convencerse de que a las operaciones de multiplicación y adición lógicas van a corresponder las conexiones en serie y en paralelo de los contactos, como se muestra en la figura 4-8.

Por medio de distintas conexiones de los contactos es posible crear esquemas que ejecuten operaciones lógicas. En la figura 4-9 se muestra un esquema de relé de contactos que realiza la operación (4-14), y en la figura 4-10 se presentan los esquemas de relé de contactos que ilustran algunas identidades de lógica matemática.

En los últimos años han alcanzado difusión los métodos de elaborar los esquemas combinatorios sin contactos que se confeccionan con ayuda de elementos semiconductores y magnéticos reunidos en esquemas integrales.

b) Composición de una fórmula lógica según una tabla dada

Al proyectar esquemas combinatorios rara vez se logra expresar directamente a modo de fórmula lógica los problemas a resolver con tal esquema. Por lo general, en la primera etapa se utiliza la descripción verbal de las tareas a solucionar con el esquema, y fundándose en ésta se consigue compilar la tabla que relaciona los valores numéricos de las variables lógicas de entrada y de salida. El paso de la tabla a la fórmula lógica es la segunda etapa de la síntesis del esquema combinatorio.

Vamos a denominar la tabla la *más sencilla*, si ésta corresponde a las operaciones de multiplicación o adición lógicas de las proposiciones iniciales. Si la tabla corresponde a la operación de multiplicación lógica, van a haber ceros en todas las líneas de la columna q y sólo estará la unidad en una sola línea correspondiente a las proposiciones iniciales verdaderas. Si la tabla corresponde a la operación de adición lógica, en todas las líneas de la columna q van a haber unidades y el cero sólo estará en la línea correspondiente a las proposiciones iniciales falsas.

Aclararemos con los ejemplos de las tablas 4-3 y 4-4 el modo de obtener la fórmula lógica para la tabla más sencilla.

En la tabla 4-3 el cero sólo está en la segunda línea de la columna q, lo que corresponde a la operación de adición lógica. Como que en esta línea $\bar{x} = 0$ e $y = 0$, entonces $q = \bar{x} + y$.

En la tabla 4-4 la unidad sólo está en la segunda línea de la columna q, lo que corresponde a la operación de multiplicación lógica. Como que en dicha línea $x = 1$ e $\bar{y} = 1$, entonces $q = x\bar{y}$.

Si la tabla no es la más sencilla, puede escribirse para ella la fórmula lógica por cualquiera de dos métodos típicos.

Primer método. En la columna q reemplacemos por ceros todas las unidades, una tras otra salvo una, compilemos las fórmulas

lógicas correspondientes a cada una de las tablas más sencillas que se obtienen así y tomemos su suma lógica.

Tabla 4-3

$q = \bar{x} + y$

x	y	q
0	0	1
1	0	0
0	1	1
1	1	1

Tabla 4-4

$q = x\bar{y}$

x	y	q
0	0	0
1	0	1
0	1	0
1	1	0

Segundo método. En la columna q reemplacemos por unidades todos los ceros, uno tras otro salvo uno, compilemos las fórmulas lógicas correspondientes a cada una de las tablas más sencillas que se obtienen así y tomemos su producto lógico.

Tabla 4-5

Adición por el módulo 2

x	y	q
0	0	0
0	1	1
1	0	1
1	1	0

Ilustremos estos métodos con el ejemplo de obtener la fórmula lógica para la tabla 4-5. Examinemos separadamente por el primer método las tablas que contienen la unidad sólo en la segunda y la tercera líneas de la columna q. A ellas corresponden las fórmulas lógicas $\bar{x}y$ y $x\bar{y}$. Tomando la suma lógica de estas fórmulas, obtenemos:

$$q = \bar{x}y + x\bar{y}.$$

Examinemos separadamente por el segundo método las tablas que contienen el cero sólo en la primera y la última líneas de la columna q. A ellas corresponden las fórmulas lógicas $x + y$ y $\bar{x} + \bar{y}$. Tomando el producto lógico de estas fórmulas, obtenemos:

$$q = (x + y)(\bar{x} + \bar{y}).$$

Como vemos, los métodos primero y segundo dan diferentes fórmulas booleanas para una misma tabla. No obstante, empleando las identidades de lógica matemática no es difícil reducir la segunda fórmula a la primera. Abriendo paréntesis y tomando en cuenta que $x\bar{x} = 0$ e $y\bar{y} = 0$, se obtiene:

$$q = x\bar{x} + x\bar{y} + y\bar{x} + y\bar{y} = x\bar{y} + y\bar{x}.$$

El método analizado es universal en el sentido de que permite obtener la fórmula lógica y, en correspondencia con ella, construir el esquema combinatorio para cualquier tabla. Sin embargo, este método, por lo general, conduce a fórmulas complejas y para reali-

zarlas se requiere gran número de elementos diversos. Por cierto, vamos a entender por *complejidad* de una fórmula el número de operaciones que constituyen esta fórmula, de manera que la complejidad de la fórmula $\bar{x}$ será el número 1 y la complejidad de la fórmula $(\bar{x}+y)(\bar{y}+z)$ será el número 5 (dos negaciones, dos adiciones, una multiplicación). A modo de ejemplo compilemos por el primer método la fórmula lógica para la tabla 4-2:

$$q=\bar{x}y\bar{z}+\bar{x}\bar{y}z+x\bar{y}z+\bar{x}yz+xyz. \qquad (4\text{-}15)$$

Esta fórmula es mucho más compleja que (4-14), aunque ambas corresponden a una misma tabla. De este modo, una importante tarea de síntesis de los esquemas combinatorios es la de simplificar o minimizar las fórmulas booleanas.

c) Simplificación de las fórmulas booleanas

Para simplificar las fórmulas booleanas se utilizan las identidades de lógica matemática examinadas en el párrafo anterior. Es cierto, que el uso con éxito de estas identidades depende de la práctica y la habilidad al manejarlas, lo que se logra después de adquirir cierta experiencia en dichas transformaciones. No obstante, existen ciertos métodos típicos que en la mayoría de los casos permiten simplificar felizmente fórmula complicada.

La simplificación de una fórmula booleana debe comenzarse hallando una de las formas siguientes: $A\bar{B}+AB$, $A+AB$, $A+\bar{A}B$, donde A y B denotan las propias variables lógicas o los productos lógicos de unas cuantas variables. Cada una de las expresiones obtenidas puede escribirse de manera más simple

$$A\bar{B}+AB=A(\bar{B}+B)=A; \qquad (4\text{-}16)$$

$$A+AB=A(1+B)=A; \qquad (4\text{-}17)$$

$$A+\bar{A}B=(A+AB)+\bar{A}B=A+B. \qquad (4\text{-}18)$$

Volvamos a la fórmula (4-15) y tratemos de simplificarla utilizando las identidades examinadas. Agrupando los sumandos primero y cuarto, así como el tercero y quinto, y aplicando las identidades (4-16) obtenemos:

$$q=\bar{x}y+\bar{x}\bar{y}z+xz.$$

La simplificación ulterior no presenta dificultades

$$q=\bar{x}(y+\bar{y}z)+xz=\bar{x}(y+z)+xz=\bar{x}y+\bar{x}z+xz=\bar{x}y+z.$$

Al simplificar las fórmulas lógicas, siempre hay que tomar en consideración la identidad $A+A=A$, de la cual se deduce que cada uno de los sumandos puede utilizarse repetidamente en combinaciones con otros sumandos. Teniendo en cuenta esta circuns-

tancia, examinemos la simplificación de la siguiente fórmula lógica:

$$q = x\bar{y}\bar{z} + xy\bar{z} + \bar{x}yz + xyz.$$

Aplicando la identidad (4-16) a los sumandos segundo y cuarto, así como al tercero y cuarto, obtenemos

$$q = x\bar{y}\bar{z} + xy + yz = x(y + \bar{y}\bar{z}) + yz = x(y + \bar{z}) + yz =$$
$$= xy + x\bar{z} + yz.$$

No se logra su simplificación ulterior con ayuda de las identidades examinadas más arriba. Surge la cuestión: ¿es acaso la forma dada la más sencilla? Esto puede verificarse probando si la misma contiene o no sumandos sobrantes. Al hacerlo vamos a denominar sobrante el sumando si en cualquiera de los juegos de variables en el que aquél se convierte en la unidad, también se vuelve igual a la unidad cualquier otro grupo de sumandos. Así pues, el sumando xy se convierte en la unidad en el juego $x = 1$, $y = 1$. En tal caso la suma de los otros dos sumandos da $x\bar{z} + yz = \bar{z} + z = 1$. De este modo, el sumando xy es sobrante y puede eliminarse de la fórmula, de lo que en definitiva se obtiene

$$q = x\bar{z} + yz.$$

Por el método de selección completa se halla la más sencilla de todas las formas que no contienen sumandos sobrantes. Pueden haber varias de estas formas.

Para aquellos casos en que el número de variables de entrada no es mayor de cuatro, la búsqueda de los tipos más sencillos de fórmulas booleanas se realiza cómodamente utilizando las tablas especiales conocidas como *mapas de Karnaugh* o *diagramas de Veitch*. La tabla 4-6 representa la carta de Karnaugh para una función booleana de tres variables.

Tabla 4-6

Aspecto del mapa de Karnaugh para la función booleana do tres variables

x y	z = 0	z = 1
0 0	1	2
0 1	3	4
1 1	5	6
1 0	7	8

Tabla 4-7

Aspecto del mapa de Karnaugh para la función booleana presentada en la tabla 4-2

x y	z = 0	z = 1
0 0		×
0 1	×	×
1 1		×
1 0		×

Cada cuadrado de dicha carta (ellos están enumerados para facilitar la explicación), corresponde a un juego de variables x, y, z. Es cómodo designar los cuadrados correspondientes a los juegos en los que la función booleana adquiere el valor 1, con crucesitas como se muestra en la tabla 4-7 que representa el mapa de Karnaugh para la función booleana dada en la tabla 4-2. La carta está formada de tal modo que las zonas correspondientes a los distintos grupos de cuadrados vecinos pueden considerarse como conjuntos de veracidad para las proposiciones correspondientes a las fórmulas booleanas más sencillas con una y dos variables (entonces debe estimarse que los bordes superior e inferior de la carta están pegados uno al otro de manera que los cuadrados superior e inferior de la misma son vecinos). Así pues, corresponden a las zonas de cuatro cuadrados vecinos las fórmulas booleanas de una variable:

$$1234 - \bar{x};\ 5678 - x;\ 3456 - y;\ 1278 - \bar{y};\ 1357 - \bar{z};\ 2468 - z.$$

Corresponden a las zonas de dos cuadrados vecinos funciones booleanas de dos variables del tipo $12 - \bar{x}\bar{y}$, $56 - xy$; $68 - xz$, etc.

Para hallar la fórmula booleana determinada por cualquier configuración de las celdas en el mapa de Karnaugh es suficiente presentar esta combinación en forma de reunión de las zonas indicadas anteriormente. La más sencilla de las representaciones de este tipo proporciona asimismo la más sencilla fórmula booleana. De este modo, para la tabla 4-7 la zona mencionada puede representarse como la reunión de la franja vertical 2468, o sea, z con la franja horizontal 34, es decir, $\bar{x}y$, lo que da $q = z + \bar{x}y$.

Para simplificar las fórmulas booleanas con un gran número de variables, se vuelve dificultoso el empleo del método descrito. En estos casos resulta útil emplear métodos especiales de minimización entre los cuales es muy efectivo el método de Quine — Mc.Closkey.

d) Ejemplos de síntesis de esquemas combinatorios

Ejemplo 4-2. Construir el esquema lógico del sumador de dígito con dos entradas (*semisumador*).

El sumador de un dígito realiza la adición de dos números de un dígito según la regla

$$0 + 0 = 0,\quad 0 + 1 = 1 + 0 = 1,\quad 1 + 1 = 10.$$

En el último caso se pone 0 en el dígito dado y se transfiere la unidad al dígito siguiente. Designemos por x e y los valores del dígito dado de los sumandos, q, el valor del dígito dado de la suma, u, p, la unidad de traspaso al dígito siguiente. Entonces el funcionamiento del semisumador será descrito mediante la tabla 4-8 en conformidad con la cual hallamos:

$$q = x\bar{y} + \bar{x}y;\ p = xy. \tag{4-19}$$

El esquema lógico del semisumador se muestra en la figura 4-11.

Ejemplo 4-3. Esquema selectivo (*descifrador*). El esquema lógico que tiene dos entradas x, y y cuatro salidas q_1, q_2, q_3, q_4, se denomina esquema selectivo si la señal 1 aparece solamente en una de las salidas para cada una de las

Tabla 4-8

Semisumador

x	y	q	p
0	0	0	0
0	1	1	0
1	0	1	0
1	1	0	1

Tabla 4-9

Descifrador

x	y	q_1	q_2	q_3	q_4
0	0	1	0	0	0
0	1	0	1	0	0
1	0	0	0	1	0
1	1	0	0	0	1

cuatro combinaciones posibles de las señales de entrada. Satisface a esta condición la tabla 4-9, según la cual las señales de salida del esquema deben ser iguales a:

$$q_1 = \bar{x}\bar{y}; \quad q_2 = \bar{x}y; \quad q_3 = x\bar{y}; \quad q_4 = xy. \qquad (4\text{-}20)$$

En la figura 4-12 se muestra la estructura lógica del esquema selectivo.

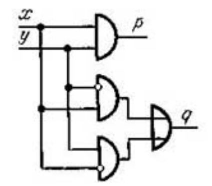


Fig. 4.11. Semisumador

Fig. 4-12. Descifrador

Ejemplo 4-4 Operador automático. Presentemos en forma muy simplificada el problema de construir el operador automático que gobierna el aterrizaje de aviones en un aeródromo. Consideremos que el aeródromo tiene una pista de aterrizaje. El avión, al aproximarse en vuelo, emite sus señales de llamada y el operador automático da el permiso para aterrizar. Si se aproximaron al aeródromo varios aviones, se da la señal de aterrizar a unos y la de mantenerse en el aire a los demás.

Asignemos a todos los aviones que tienen su base en el aeródromo los números del 1 al n. Si varios aviones llegaron simultáneamente, se da preferencia en el aterrizaje al avión con el número de orden menor. Se considera imposible y se excluye del análisis el caso en que se aproximan al aeródromo más de tres aviones. Es necesario construir el esquema combinatorio que ejecute las funciones del operador automático.

Formulemos el problema en términos de lógica matemática. Designemos por x_k la señal de llegada del k-ésimo avión. El esquema lógico combinatorio debe transmitir al k-ésimo avión una de las tres señales siguientes:

q_k^0, aterrizar inmediatamente;

q_k', dar una vuelta sobre el aeródromo;

q_k'', dar dos vueltas sobre el aeródromo.

En el caso dado es embarazoso describir el funcionamiento del esquema combinatorio con ayuda de una tabla y por eso vamos a resolver el problema de otra manera. Notemos que las señales q_k^0, q_k', q_k'' sólo pueden ser iguales a

la unidad si el k-ésimo avión se aproxima en vuelo al aeródromo, es decir, si $x_k = 1$. Por tanto, la señal x_k debe entrar como factor en las fórmulas lógicas para q_k^0, q_k', q_k'' que de esta manera pueden representarse en la forma

$$q_k^0 = A_k^0 x_k, \quad q_k' = A_k' x_k, \quad q_k'' = A_k'' x_k. \tag{4-21}$$

Las fórmulas lógicas para A_k^0, A_k' y A_k'' se obtienen pasando de k a $k+1$.

La condición $A_{k+1}^0 = 1$ se cumple si no hay ninguna señal del total $x_1, \ldots, \ldots, x_k$, es decir, si no hay señal alguna del total $x_1, \ldots, x_{k-1}$ y no hay la señal x_k. De este modo

$$A_{k+1}^0 = A_k^0 \bar{x}_k. \tag{4-22}$$

La condición $A_{k+1}^0 = 1$ se cumple si hay una señal del total $x_1, \ldots, x_k$, o sea, si hay una señal del total $x_1, \ldots, x_{k-1}$ y no hay la señal x_k, o no hay

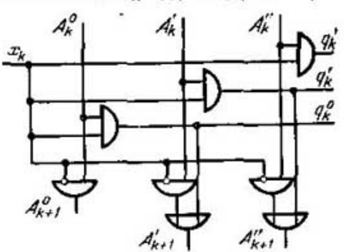


Fig. 4-13. Despachador automático

ninguna señal del total $x_1, \ldots, x_{k-1}$ y hay la señal x_k. De esta suerte.

$$A_{k+1}' = A_k' \bar{x}_k + A_k^0 x_k = A_k' \bar{x}_k + q_k^0. \tag{4-23}$$

La condición $A_{k+1}'' = 1$ se cumple si hay dos señales del total $x_1, \ldots, x^k$ es decir, si hay dos señales del total $x_1, \ldots, x_{k-1}$ y no hay x_k o si hay una señal de dicho total $x_1, \ldots, x_{k-1}$ y hay x_k. De este modo,

$$A_{k+1}'' = A_k'' \bar{x}_k + A_k' x_k = A_k'' \bar{x}_k + q_k'. \tag{4-24}$$

Las fórmulas de (4-21) a (4-24) describen por completo la estructura lógica de un elemento del esquema combinatorio que gobierna el aterrrizaje del k-ésimo avión y cuyo esquema se muestra en la figura 4-13. Construyendo dichos esquemas para cada elemento y conectándolos entre sí, obtenemos el esquema completo del operador automático.

4-4. CONCEPTO DE DISPOSITIVOS AUTOMATICOS TERMINALES

a) Esquema combinatorio como dispositivo automático terminal sin memoria

El término "dispositivo automático terminal" se utiliza para designar una clase de sistemas dinámicos numéricos que se emplean en automática, telemecánica y técnica de cálculo.

El tipo más sencillo de dispositivo automático terminal son los esquemas combinatorios ya examinados a los que, no obstante, es conveniente dar una descripción más general.

Supongamos que tenemos unas cuantas variables lógicas $u_1, \ldots, u_r$ el total de las cuales puede considerarse como la magnitud multidimensional $u = (u_1, \ldots, u_r)$. Sea f una función lógica de estas variables, que constituye cierta proposición lógica compleja

$$q = f(u_1, \ldots, u_r). \tag{4-25}$$

La variable lógica q puede considerarse como magnitud de salida del esquema combinatorio que realiza la operación lógica f con las variables $u_1, \ldots, u_r$.

Los esquemas combinatorios que encontramos en la práctica no disponen comúnmente de una salida q sino de varias, por ejemplo, $q_1, \ldots, q_m$, cada salida realiza su función lógica dependiendo de las mismas variables lógicas

$$\left.\begin{array}{l} q_1 = f_1(u_1, \ldots, u_r); \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ q_m = f_m(u_1, \ldots, u_r). \end{array}\right\} \tag{4-26}$$

La estructura de dicho esquema combinatorio se presenta en la figura 4-14.

La relación (4-26) puede escribirse más compactamente si se incluyen en el análisis la variable lógica multidimensional de salida $q = (q_1, \ldots, q_m)$ y la función lógica *multidimensional* $f = (f_1, \ldots, f_m)$. Entonces, la totalidad de las funciones (4-26) puede escribirse a modo de función multidimensional

$$q = f(u). \tag{4-27}$$

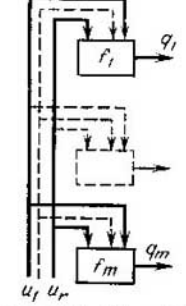


Fig. 4-14. Dispositivo automático terminal sin memoria

En lo sucesivo vamos a suponer que el esquema combinatorio efectúa instantáneamente la operación lógica, es decir, si en determinado instante ha llegado al esquema cierto valor de la magnitud de entrada $u \in U$, el valor correspondiente $q = f(u)$ se obtiene a la salida en ese mismo instante.

Para describir el funcionamiento del esquema combinatorio en el tiempo conviene introducir el concepto de *tiempo discreto*. En el eje del tiempo marcamos los instantes en los cuales puede sufrir cambios la magnitud de entrada, designándolos por $t_0, t_1, t_2 \ldots$ En lo sucesivo será conveniente designar los instantes marcados, simplemente en forma de números enteros no negativos 0, 1, 2 ... y llamarlos tiempos. La escala así obtenida se presenta en la figura 4-15 y será la escala del tiempo discreto.

Designemos por la letra n el tiempo corriente, es decir, el tiempo correspondiente al momento actual. Vamos a designar por

$u[n]$ y $q[n]$ los valores de las magnitudes de entrada y salida en el instante actual. Los valores que han obtenido las magnitudes de entrada y salida k tiempos antes del momento actual van a designarse por $u[n-k]$ y $q[(n-k)]$. La ecuación que describe el funcionamiento del esquema combinatorio en el tiempo puede escribirse ahora en la forma

$$q[n] = f(u[n]). \tag{4-28}$$

Los esquemas combinatorios son la forma más sencilla de los dispositivos automáticos terminales denominados *dispositivos automáticos terminales sin memoria*. Esta denominación pone de manifiesto que la magnitud de salida del esquema combinatorio en cada



Fig. 4-15. Escala de tiempo discreto

tiempo sólo depende de las señales de entrada en dicho tiempo y no del estado del esquema ni de las señales de entrada en los tiempos precedentes.

b) Dispositivos automáticos terminales de tipo general

Los dispositivos automáticos terminales sin memoria son el tipo más sencillo de dispositivos automáticos, siendo relativamente limitada la posibilidad de utilizarlos en calidad de elementos de dispositivos calculadores y sistemas automáticos. Se logra una considerable ampliación de las posibilidades de los dispositivos automáticos, introduciendo en su constitución elementos adicionales

$x[n]$ D $x[n-1]$

Fig. 4-16. Elemento de memoria

que retienen durante un lapso de tiempo las señales que llegan a su entrada y que reciben el nombre de *elementos de memoria*. En la figura 4-16 se ofrece la representación esquemática de un elemento de memoria. La reunión de elementos de memoria con esquemas combinatorios permite crear dispositivos automáticos terminales de tipo general.

Para el dispositivo automático terminal es característico que su estado va a determinarse no sólo por el tipo de señales de entrada en el tiempo considerado, sino también por el estado en el que se hallaba en el tiempo anterior, siendo necesario por lo general di-

ferenciar entre sí el estado del dispositivo automático y su señal de salida.

Puede servir de ejemplo de dispositivo automático terminal con memoria el esquema representado en la figura 4-17. Este consta de tres esquemas combinatorios que realizan las funciones lógicas f_1, f_2 y f_3 y de dos elementos de memoria. La particularidad de dicho esquema consiste en que las señales x_1 y x_2, siendo señales de salida de dos esquemas combinatorios, llegan a la entrada de aquellos mismos esquemas con retardo en un tiempo. Estas señales circulan dentro del esquema del dispositivo automático terminal y pueden denominarse variables que describen el estado del dispositivo automático. La señal de salida q constituye cierta función de las variables del estado.

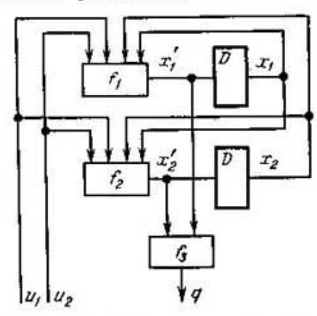


Fig. 4-17. Dispositivo automático terminal con memoria

Describamos el funcionamiento en el n-esimo tiempo del esquema examinado. Para simplificar la notación no vamos a señalar el número del tiempo n. Marcaremos con una raya los valores de las variables del estado en la entrada del elemento de memoria en el tiempo considerado. Para las señales x'_1, x'_2 y q obtendremos las ecuaciones

$$\left.\begin{aligned} x'_1 &= f_1(x_1,\ x_2,\ u_1,\ u_2); \\ x'_2 &= f_2(x_1,\ x_2,\ u_1,\ u_2); \\ q &= f_3(x'_1,\ x'_2) = \varphi(x_1,\ x_2,\ u_1,\ u_2). \end{aligned}\right\} \qquad (4\text{-}29)$$

Para simplificar la notación introduzcamos en el análisis las magnitudes multidimensionales $u = (u_1, u_2)$, $x = (x_1, x_2)$ y la función multidimensional $f = (f_1, f_2)$. Entonces las ecuaciones (4-29) se escriben en la forma

$$x' = f(x, u); \quad q = \varphi(x, u). \qquad (4\text{-}30)$$

La función $f(x, u)$ que define un nuevo estado interior del dispositivo automático terminal se denomina *función de transiciones*, y la función $\varphi(x, u)$ que determina un nuevo valor de la señal de salida se llama *función de salidas*.

Si se designan por U el conjunto de señales de entrada, Q, el de señales de salida y X, el de estados, puede darse la siguiente definición formal de dispositivo automático terminal. Se denomina *dispositivo automático terminal* al total de las cinco magnitudes

$$A = (U,\ Q,\ X,\ f,\ \varphi), \qquad (4\text{-}31)$$

donde $f: X \times U \rightarrow X$ es la función de transiciones y $\varphi: X \times U \rightarrow Q$ la función de salidas.

No se acostumbra a presentar los dispositivos automáticos terminales por medio de ecuaciones (4-30), sino mediante dos tablas

Tabla 4-10

Tabla de transiciones del dispositivo automático terminal

u	x		
	1	2	3
a	2	3	3
b	3	2	2

Tabla 4-11

Tabla de salidas del dispositivo automático terminal

u	x		
	1	2	3
a	c	c	d
b	d	c	c

denominadas, respectivamente, tabla de transiciones y tabla de salidas. Las líneas de ambas tablas corresponden a diferentes señales de entrada del dispositivo automático terminal y las columnas, a sus diferentes estados. En la intersección de líneas y columnas en la tabla de transiciones están los valores de la función $f(x, u)$ y en la tabla de salidas, los valores de la función $\varphi(x, u)$. Las tablas 4-10 y 4-11 pueden servir de ejemplo de presentación de un dispositivo automático que tiene dos señales de entrada posibles a y b, dos señales de salida posibles c y d, y tres estados posibles 1, 2, 3.

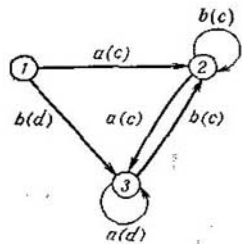


Fig. 4-18. Grafo de transiciones del dispositivo automático terminal

Otro método de preasignación de los dispositivos automáticos terminales que brinda una gran claridad es la preasignación de los dispositivos automáticos con ayuda de grafos direccionales. Los vértices del grafo representados en forma de círculos se identifican con los diversos estados del dispositivo automático. Los dos vértices x_i y x_k se conectan por el arco que va de x_i a x_k en el caso en que exista la señal u que hace pasar el dispositivo del estado x_i al x_k. Junto a cada arco se pone la señal u que realiza la transición dada y, por lo común, entre paréntesis, la señal de salida q obtenida de este modo. En la figura 4-18 se muestra el grafo correspondiente a las tablas 4-10 y 4-11.

Los dispositivos automáticos terminales son el modelo matemático de una amplia categoría de sistemas dinámicos que funcionan en el tiempo discreto y entre los cuales se encuentran también ciertos tipos de sistemas dinámicos paso a paso múltiples que se estudian en la segunda parte del libro. En la presente obra no es

posible detenerse detalladamente en la teoría de los dispositivos automáticos terminales. Por eso nos limitaremos solamente al análisis de algunos ejemplos.

Ejemplo 4-5. Generador de unidades. En la figura 4-19, *a* se muestra el grafo de un dispositivo automático terminal que en cada tiempo va a suministrar la señal 1, independientemente de la señal en la entrada, después de que a su entrada llega una vez la señal 1. La función de transiciones $x' = f(x, u)$ y la función de salidas $q = \varphi(x, u)$ de este dispositivo se muestran en la tabla 4-12 por la que hallamos:

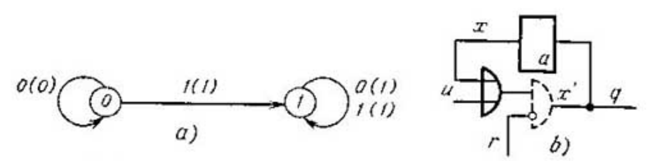


Fig. 4-19. Grafo de transiciones y estructura del generador de unidades

$$x' = q = u + x. \quad (4\text{-}32)$$

La estructura del dispositivo automático se muestra en la figura 4-19, *b* donde está prevista la existencia de la entrada auxiliar *r*, al suministrarle una señal puede suspenderse la generación de unidades.

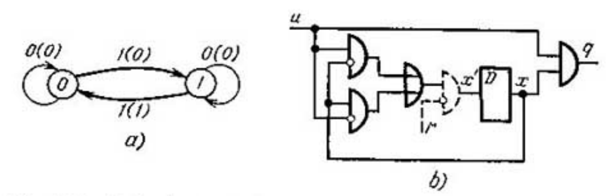


Fig. 4-20. Grafo de transiciones y estructura del generador binario

Ejemplo 4-6. Contador. En la figura 4-20, *a* está representado el grafo de un dispositivo automático terminal que da una unidad en la salida después de cada par de unidades en la entrada que se alternan con cualquier número de

Tabla 4-12

Generador de unidades

u	x	x'	q
0	0	0	0
1	0	1	1
0	1	1	1
1	1	1	1

Tabla 4-13

Contador binario

u	x	x'	q
0	0	0	0
1	0	1	0
0	1	1	0
1	1	0	1

ceros. Semejante dispositivo se denomina contador de 2. La función de transiciones y la función de salidas de este dispositivo automático se presentan en la tabla 4-13 por la que hallamos:

$$x' = u\bar{x} + \bar{u}x; \quad q = ux. \tag{4-33}$$

La estructura del contador de 2 se muestra en la figura 4-20, *b* donde está prevista la existencia de la entrada auxiliar *r* suministrando una señal a la cual puede cesar el funcionamiento del contador y pasarlo al estado inicial 0.

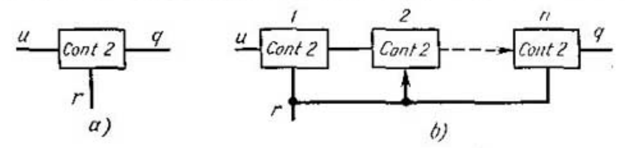


Fig. 4-21. Contador binario y contador de 2^n

La representación convencional del contador de 2 se muestra en la figura 4-21, *a*. En la fagura 4-21, *b* se presenta el esquema de *n* contadores de 2 conectados en serie y formando un contador de 2^n.

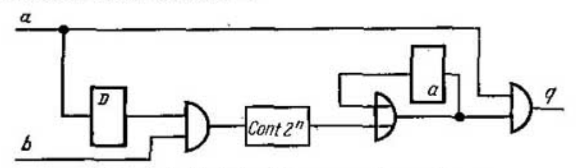


Fig. 4-22. Modelo del proceso de enseñanza

Ejemplo 4-7. Modelo del proceso de enseñanza. Basándose en el contador de 2^n puede construirse un dispositivo que simula, aunque de forma muy simplificada, el proceso de enseñanza. El esquema de tal dispositivo se muestra en la figura 4-22. El mismo tiene dos entradas *a* y *b* y se enseña a que después de

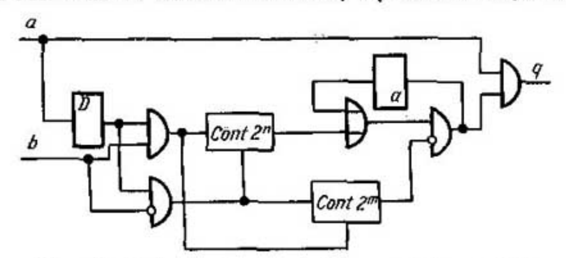


Fig. 4-23. Modelo perfeccionado del proceso de enseñanza

la unidad en la salida *a* va a seguir la unidad en la salida *b*. Si esto sucede 2^n veces (no necesariamente de forma sucesiva y posiblemente con muchas alteraciones), el dispositivo aprende a prever la unidad en la salida *b* a continuación de la llegada de la unidad a la entrada *a* y lo expresa presentando la unidad en la salida *q* cada vez que aparece la unidad en la entrada *a*.

El esquema en la figura 4-22 simula un proceso de enseñanza muy primitivo. Valiéndonos de ciertos ingenios el dispositivo puede perfeccionarse de tal modo que solamente aprenda a saber que "detrás de a sigue b", en caso de que este suceso haya ocurrido 2^n veces seguidas sin alteraciones y que sea capaz de "olvidarlo" si aparecen 2^m sucesos "detrás de a no sigue b", en caso de que esta sucesión no se interrumpa por un suceso "detrás de a sigue b". El esquema de tal dispositivo se presenta en la figura 4-23.

PROBLEMAS PARA EL CAPITULO 4

4-1. ¿Con qué operación lógica puede describirse el funcionamiento del esquema de mando en el ejemplo 4-1?

4-2. ¿En qué consiste la diferencia entre una fórmula booleana y una función booleana?

4-3. Comprobar con ayuda de los esquemas de conmutación si se cumplen las identidades $x + \bar{x} = 1$, $x + 1 = 1$, $x \cdot 0 = 0$, $xx = x$.

4-4. Componer el circuito de conexiones del alumbrado para un local que tiene dos puertas en los extremos opuestos de modo que pueda encenderse la luz al entrar por cualquier puerta, y apagarla al salir por cualquiera de ellas.

4-5. Simplificar las siguentes fórmulas booleanas:

$$xyz + x\bar{y}z + xy\bar{z}v;$$

$$\bar{x}y\bar{z} + x\bar{y}\bar{z} + xy\bar{z};$$

$$xy + z + (\overline{xy + z})(zv + x).$$

4-6. Construir el esquema lógico del sumador binario de un dígito con tres entradas designando por x e y los valores de los dígitos de los sumandos; z, la unidad de pase del dígito precedente de la suma; p, la unidad de pase al dígito siguiente y q, el valor del digito dado de la suma.

4-7. Construir el esquema lógico del sumador binario de un dígito con tres entradas utilizando sumadores binarios con dos entradas.

4-8. Construir el esquema lógico del sumador binario de dígitos múltiples utilizando los resultados de los problemas 4-6 y 4-7.

Capítulo quinto

CORRELACIONES DE LA TEORIA DE LOS CONJUNTOS EN LA TEORIA DE PROBABILIDADES Y ELEMENTOS DE ESTADISTICA MATEMATICA

5-1. CONCEPTO DE PROBABILIDAD

a) Suceso y espacio de resultados de experimento

La base de la teoría de probabilidades es el concepto de *suceso aleatorio*. Relacionamos este suceso con la realización de cierto experimento.

Al realizar dicho experimento nos interesan sus consecuencias, o, como se dice en la teoría de probabilidades, su *resultado*. Vamos a suponer que con el total de condiciones dado las consecuencias del experimento pueden ser un número finito de resultados $z_1, \ldots, z_m$, cuyo total designaremos por

$$Z = \{z_1, \ldots, z_m\}. \tag{5-1}$$

Es evidente que a consecuencia de cada experimento tiene lugar sin falta algún resultado $z \in Z$ y no puede haber experimento a consecuencia del cual haya dos o más resultados. Vamos a denominar al conjunto Z *espacio de resultados* del experimento y los elementos $z \in Z$, *sucesos elementales* en el espacio Z.

Sin embargo en la práctica presentan el máximo interés no los propios sucesos elementales sino cierto grupo de ellos que son subconjuntos del conjunto Z. Vamos a llamar *suceso* en el espacio de resultados del experimento Z a cualquier subconjunto S del conjunto Z:

$$S \subseteq Z. \tag{5-2}$$

Cuando decimos que *occurre* o *se realiza* el suceso S, sobreentendemos que el suceso elemental z, que es el resultado del experimento, está contenido en S.

Para sucesos cualesquiera S_1 y S_2 pertenecientes al espacio de resultados del experimento Z son válidas las siguientes definiciones:

Se denomina *reunión* $S_1 \cup S_2$ de los sucesos S_1 y S_2 al suceso que consiste en la realización siquiera de uno de los sucesos S_1 y S_2.

Se llama *conjunción* $S_1 \cap S_2$ de los sucesos S_1 y S_2 al suceso que consiste en la realización de S_1 y S_2. Los sucesos S_1 y S_2 se denominan incompatibles si la realización de uno de ellos excluye la posibilidad de que se efectúe el otro, es decir, si $S_1 \cap S_2 = \varnothing$.

Se llama *complemento* $\bar{S}$ del suceso S al suceso que consiste en el incumplimiento del suceso S.

El *suceso cierto* consiste en la realización siquiera de uno de los sucesos del espacio Z.

El *suceso imposible* es el conjunto vacío $\varnothing$ consistente en que no se realiza ninguno de los sucesos del espacio Z.

Ejemplo 5-1. El experimento consiste en que se tiran dos dados de juego cada uno de los cuales puede caer indicando de uno a seis tantos. Introduzcamos en el análisis el conjunto $I = \{1, 2, \ldots, 6\}$. El resultado del experimento será el par ordenado $z = (i, k)$, donde $i, k \in I$, por cierto, i es el resultado al tirar el primer dado y k, el resultado al tirar el segundo. Es cómodo representar los resultados del experimento mediante los puntos en el plano (i, k) como se muestra en la figura 5-1. Como vemos, el espacio de resultados del experimento constará de 36 puntos $Z = \{(1,1), (1,2), \ldots, (6,6)\}$. Aquí los sucesos pueden ser muy variados. Designemos por S_j el suceso que consiste en que el dado cae señalando j tantos. Entonces $S_2 = \{(1,1)\}$; $S_7 = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}$; $S_{11} = \{(6,5), (5,6)\}$, $S_{12} = \{(6,6)\}$, etc.

Fig. 5-1. Espacio de resultados del experimento al tirar dos dados de juegos

Los sucesos que pueden tener lugar en el espacio finito de resultados de experimento Z con el conjunto vacío $\varnothing$ unido a él, forman la clase de sucesos $\mathfrak{F}$ denominada *clase finita aditiva, campo booleano* o *álgebra booleana* y satisface las condiciones siguentes:

si $S \in \mathfrak{F}$, entonces $\bar{S} \in \mathfrak{F}$;

si $S_1 \in \mathfrak{F}$ y $S_2 \in \mathfrak{F}$, entonces $S_1 \cup S_2 \in \mathfrak{F}$.

Empleando sucesivamente estas condiciones puede demostrarse que pertenecerán al campo booleano $\mathfrak{F}$ la reunión o intersección de cualquier número finito de conjuntos S_k de $\mathfrak{F}$, así como las diferencias de estos conjuntos.

Si el espacio Z contiene un número finito de elementos, la clase de todos los sucesos posibles en dicho espacio será también finita. En este caso es conveniente presentar cierta clase de conjuntos $\mathfrak{F}_0$ en el espacio Z que puede denominarse clase inicial. Si a los conjuntos de $\mathfrak{F}_0$ se aplican un número finito de veces las condiciones indicadas, puede obtenerse la clase completa de sucesos que obedecen a esta condición, es decir, el campo booleano $\mathfrak{F}$. En tal caso se dice que este campo $\mathfrak{F}$ ha sido originado por la clase inicial $\mathfrak{F}_0$. Existen diversos métodos para elegir la clase inicial $\mathfrak{F}_0$ cada uno de los cuales proporciona su campo booleano.

Ejemplo 5-2. Si $\mathfrak{F}_0 = Z$, entonces $Z \in \mathfrak{F}$, Luego, $\bar{Z} = \varnothing \in \mathfrak{F}$. Por cuanto $Z \cup \varnothing = Z$ y $Z \cap \varnothing = \varnothing$, entonces $\mathfrak{F} = \{Z, \varnothing\}$ es un campo booleano.

Ejemplo 5-3. Supongamos que $\mathfrak{F}_0$ consta de todos los sucesos elementales en Z. En este caso el campo booleano $\mathfrak{F}$ contiene un conjunto vacío, todos los conjuntos monoelementales de Z, todos los conjuntos bielementales de Z, ..., y por último, el propio conjunto Z.

b) Concepto de probabilidad

En la teoría de las probabilidades cada resultado de un experimento $z \in Z$ se considera como una magnitud aleatoria. Esto significa que no se conoce por anticipado cuál de los resultados va a

tener lugar al realizar el experimento. Sin embargo es evidente, en primero, que a resultas de cada experimento debe tener lugar sin falta algún resultado $z \in Z$ y, en segundo, que como consecuencia de un experimento no puede haber dos resultados.

Intuitivamente se entiende por *probabilidad* de un resultado la medida numérica que caracteriza la posibilidad objetiva de un resultado dado de un experimento. Esto significa que se atribuye a cada uno de los elementos $z \in Z$ cierto peso p, es decir, un número real positivo al cual se han impuesto ciertas limitaciones, a saber:

1) p es tanto mayor cuanto mayor es la seguridad de que va a tener lugar precisamente el resultado dado del experimento $z \in Z$;

2) la suma de los pesos p para todos los resultados posibles del experimento debe ser igual a la unidad.

La asignación del peso $p = 0$ a cualquier elemento $z \in Z$ significa que el resultado dado del experimento es imposible, o sea, constituye un *suceso imposible*. Si a algún elemento $z \in Z$ se atribuye un peso $p = 1$, el resultado dado constituye un *suceso cierto*. Efectivamente, en virtud de la segunda limitación deben asignarse pesos $p = 0$ a todos los elementos restantes del conjunto Z, es decir, los sucesos correspondientes a estos elementos son imposibles. Pero, como que obligatoriamente debe tener lugar algún resultado del experimento, será precisamente el resultado a que se asigna el peso $p = 1$.

El total de los pesos que se atribuyen a los distintos elementos del conjunto Z representa cierto conjunto P con elementos p, y el proceso propiamente dicho de asignación de pesos constituye el reflejo del conjunto Z en el conjunto P, o sea, determina las probabilidades $p \in P$ como funciones del resultado del experimento $z \in Z$, lo que puede escribirse en la forma

$$p = p(z). \tag{5-3}$$

Conforme a las limitaciones expuestas anteriormente las magnitudes $p(z)$ representan números reales que obedecen a las condiciones

$$p(z) \geqslant 0; \qquad \sum_{z \in Z} p(z) = 1. \tag{5-4}$$

El total de los pesos $p(z)$ para todas las $z \in Z$ se llama *distribución de probabilidades* en el espacio de resultados del experimento Z y cada peso $p \in P$ atribuido a un resultado elemental $z \in Z$ se denomina probabilidad de un resultado dado.

c) Probabilidad de un suceso aleatorio

Hemos definido el suceso aleatorio S como cierto subconjunto del conjunto de resultados del experimento Z. Por tanto, un suceso aleatorio representa un conjunto que consta de una parte de los

elementos del conjunto Z. Entonces, se entiende por probabilidad del suceso S designada por P_S o $P(S)$, la suma de los pesos que constituyen este suceso de los elementos:

$$P(S) = P_S = \sum_{z \in S} p(z). \tag{5-5}$$

Son casos particulares de esta fórmula:

1) el caso en que S no contiene ningún elemento de Z, o sea, es un conjunto vacío; según (5-5), la probabilidad de un conjunto vacío es igual a cero

$$P_{\varnothing} = 0; \tag{5-6}$$

2) el caso en que S coincide con Z, es decir, contiene todos los resultados posibles del experimento. Como que no puede haber ningún resultado de este último que no forme parte de Z, puede considerarse todo el espacio Z como un conjunto universal y asignársele un peso igual a la unidad. Esto concuerda absolutamente con la fórmula (5-5), la que, teniendo en cuenta (5-4), da:

$$P_Z = \sum_{z \in Z} p(z) = 1. \tag{5-7}$$

d) Espacio probabilístico

Los razonamientos precedentes respecto a la definición de probabilidades gozan de un carácter algo intuitivo y no garantizan la posibilidad de calcular las probabilidades para cualquier grupo de sucesos capaces de ocurrir en la clase de sucesos $\mathfrak{F}$. Para resolver este problema hay que analizar la distribución de probabilidades no en el espacio de resultados del experimento Z, sino en toda la clase de sucesos $\mathfrak{F}$ relacionados con dicho espacio.

Se denomina *medida probabilística* en la clase de sucesos $\mathfrak{F}$ que representan un campo booleano, la función $P(S)$ que satisface las condiciones siguientes:

1) $P(S) \geqslant 0$, para cualquier $S \in \mathfrak{F}$;

2) si $S_1, S_2, \ldots, S_n$ es la secuencia de sucesos incompatibles por pares de $\mathfrak{F}$, entonces $P\left(\bigcup_{k=1}^{n} S_k\right) = \sum_{k=1}^{n} P(S_k)$;

3) $P(Z) = 1$.

Se denomina campo de probabilidad el par $(\mathfrak{F}, P)$, es decir, el campo de conjuntos y la medida probabilística en éste.

Se llama espacio probabilístico a la tríada $(Z, \mathfrak{F}, P)$, es decir, al espacio de sucesos elementales, al campo de conjuntos presentado en dicho espacio y a la medida probabilística definida en el campo de conjuntos.

5-2. CALCULO DE PROBABILIDADES

a) Métodos de atribución de la medida probabilística

Hay una serie de métodos para atribuir pesos a ciertos elementos z del espacio Z. Cuál de los métodos se ha de utilizar en uno u otro caso depende del carácter del experimento a realizar y de la información de que disponemos. Examinemos los métodos principales.

1. *Determinación de la probabilidad por medio de frecuencia.* Analicemos el experimento con espacio de resultados $Z = \{z_1, \ldots, z_m\}$, que podemos repetir muchas veces en iguales condiciones. Supongamos que se efectuaron N experimentos durante los cuales se obtuvo N_z veces el resultado $z \in Z$ que nos interesa. Se denomina frecuencia del resultado z al número relativo de casos en los cuales tuvo lugar el resultado deseado z, es decir, la magnitud

$$q(z) = \frac{N_z}{N}. \tag{5-8}$$

No es difícil cerciorarse de que las frecuencias $q(z)$ de todos los resultados posibles $z \in Z$ obedecen a las condiciones (5-4).

Con un pequeño número de experimentos la frecuencia tiene, en notable grado, carácter aleatorio. Así pues, en caso de tirar 10 veces una moneda ésta puede caer 2 veces de cara y al lanzarla otras 10 veces puede caer, por ejemplo, 8 veces. No obstante, la práctica demuestra que al aumentar el número de experimentos la frecuencia de los diferentes resultados pierde en considerable medida su carácter aleatorio y tiene la tendencia de aproximarse con pequeñas fluctuaciones a cierto valor medio que puede considerarse precisamente como la probabilidad del suceso. Sin embargo, esta aproximación de la frecuencia a la probabilidad, al aumentar el número de experimentos, difiere de la tendencia a un límite en sentido matemático.

Así pues, no es imposible que al lanzar 10 veces la moneda ésta caiga de cara todas las 10 veces. Por cuanto el resultado de lanzar cada moneda no depende de los resultados de los lanzamientos anteriores no es nada imposible que caiga de cara en 1000 ocasiones al lanzar la moneda 1000 veces. Pero las probabilidades de tal suceso son tan pequeñas que puede considerarse prácticamente irrealizable.

En general, al aumentar el número de experimentos la frecuencia se aproxima a la probabilidad en el sentido de que la probabilidad de desviaciones de alguna consideración de la frecuencia respecto a la probabilidad llega a ser tan pequeña que es despreciable. Por consiguiente, si hay la posibilidad de repetir el experimento varias veces en idénticas condiciones, las frecuencias de algunos

resultados $q(z)$ pueden tomarse como las probabilidades de estos resultados. La cuestión de cuántos experimentos deben considerarse suficientes al determinar las probabilidades por medio de la frecuencia y cuál es el grado de certeza de los resultados obtenidos al hacerlo, será estudiada más detalladamente en los capítulos dedicados a los métodos de estadística matemática.

No obstante, las dificultades debidas a la múltiple repetición de los experimentos en condiciones reales obligan a utilizar otros métodos de determinación de la probabilidad.

2. Uno de los métodos que puede usarse con bastante frecuencia (aunque no siempre) es el método basado en el *principio de iguales posibilidades.*

El principio de iguales posibilidades se emplea cuando no tenemos fundamento para dar preferencia a cualquier resultado del experimento respecto a otros resultados. En este caso hay que considerar que existen iguales posibilidades para cualquier resultado del experimento y atribuirles iguales probabilidades. Si en este caso el espacio de resultados del experimento Z consta de m elementos y el suceso S contiene r elementos de Z, la probabilidad de este suceso será igual a:

$$P(S)=\frac{r}{m}. \qquad (5\text{-}9)$$

3. El método más extendido para atribuir la medida probabilística es la determinación de las llamadas *probabilidades apriorísticas.* Las probabilidades apriorísticas se determinan recopilando durante un prolongado lapso de tiempo los datos estadísticos sobre el suceso o fenómeno que nos interesan.

Habitualmente, en las diversas esferas de la ciencia, técnica y vida social se efectúa la recopilación de datos estadísticos sobre unos u otros sucesos o fenómenos que pertenecen a la categoría de los aleatorios, es decir, cuyas leyes no pueden describirse con rigor matemático y cuya aparición no puede pronosticarse de manera fidedigna. El estudio de cuán frecuentemente ha ocurrido suceso en el pasado permite determinar la probabilidad de este suceso y, por tanto, pronosticar con determinado grado de certeza la aparición de dicho suceso en el futuro.

4. Las probabilidades apriorísticas sólo proporcionan una descripción fidedigna de los fenómenos en caso de mantenerse en el presente las condiciones en las cuales ha sucedido el fenómeno en el pasado. Sin embargo, está muy lejos de suceder siempre así. Por eso tiene gran importancia precisar las condiciones reales existentes en la actualidad, lo que puede hacerse realizando un experimento especialmente preparado. Las probabilidades determinadas fundándose tanto en los datos estadísticos del pasado como en un experimento preparado "ad hoc" se denominan *probabilidades aposterioristicas.*

b) Propiedades de la medida probabilística

Ciertas propiedades de la medida probabilística son muy útiles para simplificar el cálculo de probabilidades de diversos sucesos. Examinemos estas propiedades para los sucesos X_1 y X_2 en el espacio de resultados de la prueba Z.

1. Si $X_1 \cap X_2 = \varnothing$, entonces

$$P(X_1 \cup X_2) = P(X_1) + P(X_2), \tag{5-10}$$

lo que se infiere directamente de la definición de medida probabilística.

2. De las condiciones $X \cup \bar{X} = Z$, $X \cap \bar{X} = \varnothing$ hallamos:

$$P(X \cup \bar{X}) = P(X) + P(\bar{X}) = 1, \tag{5-11}$$

de donde

$$P(\bar{X}) = 1 - P(X). \tag{5-12}$$

3. Sea $X_1 \subseteq X_2$. Entonces X_2 puede representarse en la forma $X_2 = X_1 \cup (X_2 \setminus X_1)$, siendo $X_1 \cap (X_2 \setminus X_1) = \varnothing$. Luego, $P(X_2) = P(X_1) + P(X_2 \setminus X_1)$. Por cuanto $X_2 \setminus X_1 \subseteq Z$, lo cual significa que $P(X_2 \setminus X_1) \geqslant 0$, obtenemos:

$$P(X_2) \geqslant P(X_1); \tag{5-13}$$

$$P(X_2 \setminus X_1) = P(X_2) - P(X_1). \tag{5-14}$$

4. Para X_1 y X_2 arbitrarias de Z representamos su reunión en forma de la reunión de dos conjuntos no intersecados

$$X_1 \cup X_2 = X_1 \cup [X_2 \setminus (X_1 \cap X_2)], \tag{5-15}$$

de suerte que

$$P(X_1 \cup X_2) = P(X_1) + P[X_2 \setminus (X_1 \cap X_2)]. \tag{5-16}$$

Tomando en consideración que $X_1 \cap X_2 \subseteq X_2$, obtenemos:

$$P(X_1 \cup X_2) = P(X_1) + P(X_2) - P(X_1 \cap X_2). \tag{5-17}$$

Ejemplo 5-4. Se escoge al azar una carta de una baraja de 52 cartas. ¿Qué probabilidad hay de que:

1) sea del palo de corazones o el rey de tréboles?
2) sea del palo de corazones o uno de los reyes?

Designemos por X el conjunto de corazones; Y, el conjunto de reyes y V, el rey de tréboles, de modo que $P(X) = 1/4$, $P(Y) = 1/13$ y $P(V) = 1/52$.

1. Calculemos $P(X \cup V)$. Por cuanto $X \cup V = \varnothing$, entonces $P(X \cup V) = P(X) + P(V) = \frac{1}{4} + \frac{1}{52} = \frac{7}{26}$.

2. Calculemos $P(X \cup Y)$. Como que $X \cup Y \neq \varnothing$, entonces $P(X \cup Y) = (P(X) + P(Y) - P(X \cap Y)$. Pero $X \cap Y$ es el rey de corazones, o sea, $P(X \cap Y) = 1/52$. Luego,

$$P(X \cup Y) = \frac{1}{4} + \frac{1}{13} - \frac{1}{52} = \frac{4}{13}.$$

5-3. PROBABILIDADES CONDICIONALES

a) Concepto de probabilidad condicional

Examinemos cierto suceso T en el espacio de resultados del experimento Z. Por definición, la probabilidad de este suceso es

$$P(T) = \sum_{z \in T} p(z). \tag{5-18}$$

Supongamos ahora que sabemos que ocurrió algún otro suceso S. Se inquiere, ¿de qué manera el saber que ocurrió el suceso S influirá sobre la probabilidad del suceso T?

La probabilidad del suceso T con la condición de que ocurrió otro suceso S se denomina *probabilidad condicional* del suceso T y se designa por $P_S(T)$ o $P(T|S)$. A diferencia de ella se llama *probabilidad incondicional* del suceso T a la probabilidad $P(T)$. Ante nosotros se plantea la tarea de hallar la correlación que vincula estas dos probabilidades.

Simplifiquemos primero el problema. Vamos a considerar que el conjunto T sólo consta de un elemento $z \in Z$. Al determinar la probabilidad incondicional $p(z)$ partimos de que puede ocurrir cualquier $z \in Z$. Si se realiza un número suficientemente grande de experimentos, entonces

$$p(z) = \frac{\text{Número de experimentos con las } z \in Z \text{ dadas}}{\text{Número total de experimentos}} = \frac{N_z}{N}. \tag{5-19}$$

Si se impone la condición de que ocurrió el suceso S, entonces, con la misma, deberán excluirse todos aquellos resultados del experimento con los que no ocurre el suceso S. Es decir, al determinar $p(z|S)$ hay que cambiar N por N_S, el número de experimentos, en los cuales ocurrió el suceso S, y N_z por $N_{(z,\,S)}$, el número de éstos en los que ocurrieron tanto la z dada como el suceso S. Pero en el caso examinado

$$(z, S) = z \cap S = \begin{cases} z, & \text{si } z \in S; \\ \varnothing, & \text{si } z \notin S. \end{cases} \tag{5-20}$$

Por eso, $N_{(z,\,S)} = N_{z \cap S}$. De esta suerte,

$$p(z|S) = \frac{N_{z \cap S}}{N_S}. \tag{5-21}$$

Dividiendo el numerador y denominador de esta expresión por el número total de experimentos N, obtenemos:

$$p(z|S) = \frac{p(z \cap S)}{P(S)}. \tag{5-22}$$

Analicemos ahora el caso en que T es un subconjunto arbitrario del conjunto Z. Teniendo en cuenta que la probabilidad de un suceso es igual a la suma de las probabilidades de los elementos que constituyen dicho suceso, obtenemos:

$$P(T \mid S) = \sum_{z \in T} p(z \mid S) = \frac{\sum_{z \in T} p(z \cap S)}{P(S)}. \qquad (5\text{-}23)$$

Examinemos la suma en el numerador de la expresión dada. Sólo forman parte de esta los sumandos para $z \in T$. Pero en ella pueden dejarse de considerar los sumandos para $z \notin S$ ya que entonces $p(z \cap S) = 0$. Por consiguiente, dicha suma consta de sumandos para los que se cumplen las condiciones $z \in T$ y $z \in S$. Es decir, la sumación puede efectuarse por $z \in T \cap S$. Pero con $z \in S$ $p(z \cap S) = p(z)$.

Por eso

$$\sum_{z \in T} p(z \cap S) = \sum_{z \in T \cap S} p(z) = P(T \cap S). \qquad (5\text{-}24)$$

Teniendo en cuenta esto obtenemos:

$$P(T \mid S) = \frac{P(T \cap S)}{P(S)}, \qquad (5\text{-}25)$$

lo que a menudo se escribe en la forma

$$P(T \cap S) = P(T \mid S)\, P(S). \qquad (5\text{-}26)$$

A veces se encuentran casos en que nuestro conocimiento de que ocurrió el suceso S no influye sobre la probabilidad del suceso T. En otras palabras, la probabilidad del suceso T no depende de si ocurrió o no el suceso S, de manera que

$$P(T \mid S) = P(T). \qquad (5\text{-}27)$$

En este caso se dice que los sucesos T y S son *independientes*. Para los sucesos independientes se cumple la correlación

$$P(T \cap S) = P(T)\, P(S), \qquad (5\text{-}28)$$

que frecuentemente se denomina fórmula de multiplicación de probabilidades.

b) Magnitudes aleatorias bidimensionales

Muy a menudo, como resultado del experimento nos interesan varias magnitudes aleatorias en lugar de una, por ejemplo, dos x e y. Así pues, al tomar una carta de la baraja pueden interesarnos su palo x y valor (número de tantos) y. Al elegir en un bosque árboles para una obra de construcción, nos interesan la altura del árbol x y su diámetro y. Al verificar el buen estado de los con-

densadores nos interesan los valores de la capacidad x y la tensión disruptiva y.

En todos estos ejemplos hay que operar con dos conjuntos de magnitudes aleatorias $X = \{x, \ldots, x_n\}$ e $Y = \{y_1, \ldots, y_m\}$. Pero estos conjuntos no son independientes, sino que están vinculados entre sí de un modo determinado; a saber, el resultado del experimento z son dos magnitudes aleatorias $x \in X$ e $y \in Y$ de modo que $z = (x, y)$. Así el conjunto de los resultados del experimento Z es el producto directo de los conjuntos X e Y:

$$Z = X \times Y. \tag{5-29}$$

Examinemos algunos sucesos y sus probabilidades que pueden interesarnos al estudiar las magnitudes aleatorias bidimensionales. Vamos a ilustrar los conceptos introducidos con el ejemplo de la altura x y el diámetro y de los árboles en el bosque.

Supongamos que $x \in X$ e $y \in Y$ son resultados de cierto experimento. Por ejemplo, $x = 30$ m e $y = 25$ cm. Entonces puede surgir la necesidad de examinar las probabilidades siguientes:

1) $p(x)$, la probabilidad de que la altura del árbol sea 30 m. En el caso dado no nos interesa el diámetro del árbol;

2) $p(y)$, la probabilidad de que el diámetro del árbol sea 25 cm. En este caso no nos interesa la altura del árbol. Las probabilidades $p(x)$ y $p(y)$ son probabilidades incondicionales de los índices x e y y proporcionan las distribuciones de probabilidades respectivamente en los conjuntos X e Y que satisfacen las condiciones (5-4).

3) $p(x, y)$, la probabilidad de que el árbol tiene una altura de 30 m y un diámetro de 25 cm. Esta es la distribución conjunta de probabilidades de dos índices x e y presentada en el espacio $Z = X \times Y$ con los elementos $z = (x, y)$ que obedece a las condiciones

$$p(x, y) \geqslant 0, \qquad \sum_{x, y} p(x, y) = 1; \tag{5.30}$$

4) por último puede encontrarse la probabilidad condicional $p(x|y)$, es decir, la probabilidad de que el árbol con diámetro de 25 cm tiene una altura de 30 m. En el caso dado los árboles de otro diámetro simplemente no se consideran.

Para relacionar entre sí todas las probabilidades examinadas, designemos por x_i cierto resultado $x \in X$, y por y_j cierto resultado $y \in Y$ analizando dos sucesos en el espacio $Z = X \times Y : T_i = \{(x_i, y_1), \ldots, (x_i, y_m)\}$ y $S_j = \{(x_1, y_j), \ldots, (x_n, y_j)\}$.

El suceso T_i consiste en que se obtuvieron el resultado dado $x_i \in X$ y cualquier $y \in Y$. Como que en cualquier experimento siempre tiene lugar algún $y \in Y$, el suceso T_i consiste en que se obtuvo el resultado x_i. De este modo, $T_i = x_i$. Análogamente, $S_j = y_j$. No es difícil ver asimismo que $T_i \cap S_j = (x_i, y_j)$. Por la

fórmula (5-25) hallamos

$$p(x_i \mid y_j) = \frac{p(x_i, y_j)}{p(y_j)}. \qquad (5\text{-}31)$$

Por cuanto x_i puede ser cualquier $x \in X$ e y_j puede ser cualquier $y \in Y$, omitiendo los subíndices obtenemos para $x \in X$ e $y \in Y$ cualesquiera:

$$p(x \mid y) = \frac{p(x, y)}{p(y)}, \qquad (5\text{-}32)$$

lo que puede escribirse en la forma

$$p(x, y) = p(x \mid y)\, p(y). \qquad (5\text{-}33)$$

Si la probabilidad del resultado x no depende de qué resultado y se obtuvo, entonces los resultados x e y son *independientes*. Para los resultados independientes se cumple la correlación

$$p(x \mid y) = p(x), \qquad (5\text{-}34)$$

de modo que la fórmula (5-33) toma la forma:

$$p(x, y) = p(x)\, p(y). \qquad (5\text{-}35)$$

Notemos más adelante que al examinar una magnitud bidimensional nos es indiferente cuál de las magnitudes x o y tomamos como primera y cuál como segunda. Intercambiando de lugares x e y en la fórmula (5-33), obtenemos:

$$p(x, y) = p(y \mid x)\, p(x). \qquad (5\text{-}36)$$

Comparando (5-33) y (5-36) hallamos que

$$p(x, y) = p(x)\, p(y \mid x) = p(y)\, p(x \mid y). \qquad (5\text{-}37)$$

c) Fórmula de la probabilidad completa

Examinemos cierto espacio de resultados del experimento Z. Sea $T \subseteq Z$ un suceso en el espacio Z, y el sistema de conjuntos $\{S_1, \ldots, S_l\}$ representa cierto fraccionamiento de dicho espacio, es decir, obedece a las condiciones

$$Z = S_1 \cup \ldots \cup S_l, \quad S_i \cap S_k = \varnothing \text{ para } i \neq k. \qquad (5\text{-}38)$$

En diversas ocasiones presenta interés expresar la probabilidad incondicional del suceso T mediante las probabilidades condicionales $P(T \mid S_1), \ldots, P(T \mid S_l)$ y las probabilidades incondicionales $P(S_1), \ldots, P(S_l)$.

Para obtener la correlación necesaria representaremos el espacio de resultados del experimento Z en la forma $Z = S_1 \cup \bar{S}_1$, donde $\bar{S}_1 = S_2 \cup \ldots \cup S_l$. Entonces

$$P(T) = P(T \mid S_1 \cup \bar{S}_1) = P[T \cap (S_1 \cup \bar{S}_1)] = P(T \cap S_1) + P(T \cap \bar{S}_1). \qquad (5\text{-}39)$$

Vamos a considerar a $\bar{S}_1$ como un nuevo espacio de resultados del experimento Z_1, es decir, $Z_1 = \bar{S}_1 = S_2 \cup \ldots \cup S_i$ o $\bar{S}_1 = S_2 \cup \cup \bar{S}_2$, donde $\bar{S}_2 = S_3 \cup \ldots \cup S_i$. Entonces

$$P(T \cap \bar{S}_1) = P[T \cap (S_2 \cup \bar{S}_2)] = P(T \cap S_2) + P(T \cap \bar{S}_2). \quad (5\text{-}40)$$

Luego,

$$P(T) = P(T \cap S_1) + P(T \cap S_2) + P(T \cap \bar{S}_2). \quad (5\text{-}41)$$

Al continuar realizando más adelante transformaciones análogas obtenemos:

$$P(T) = \sum_i P(T \cap S_i) = \sum_i P(T \mid S_i) P(S_i). \quad (5\text{-}42)$$

La fórmula (5-42) recibió el nombre de fórmula de la probabilidad completa.

5-4. MAGNITUDES ALEATORIAS CONTINUAS Y SUS DISTRIBUCIONES

a) Concepto de magnitud aleatoria continua

Hasta aquí hemos supuesto que el espacio de resultados del experimento Z constituye un conjunto finito. Sin embargo, en muchos casos el resultado z del experimento puede ser cualquier valor en cierto intervalo $a \leqslant z \leqslant b$. Entonces z es una magnitud aleatoria continua en que el espacio de resultados del experimento será todo el intervalo de sus valores posibles:

$$Z = \{z : a \leqslant z \leqslant b\}, \quad (5\text{-}43)$$

que contiene un número infinito de puntos. En este caso ya no es posible hablar de la probabilidad de un resultado particular, puesto que con un número infinitamente grande de resultados posibles el peso de cada resultado será igual a cero. Por eso para una magnitud aleatoria continua la distribución de probabilidades se determina de diferente manera que en el caso discreto. Al hacerlo, se utilizan dos tipos de distribución denominados función de distribución de probabilidades y densidad de distribución de probabilidades.

b) Función de distribución de probabilidades

Sean Z, el espacio de resultados del experimento de una magnitud aleatoria unidimensional que toma valores reales, y R, el conjunto de todos los números reales cuyos elementos designemos por x. La función de distribución de probabilidades $F(x)$ determina la probabilidad de que los valores de la magnitud aleatoria z no excedan de la x dada:

$$F(x) = P(-\infty < z \leqslant x). \quad (5\text{-}44)$$

Puede representarse el carácter general de la función $F(x)$ basándose en algunas de sus propiedades:

1) Por cuanto, para $x_2 > x_1$, el intervalo $(-\infty, x_2]$ está totalmente contenido en el intervalo $(-\infty, x_1]$, según (5-44) $F(x_2) \geqslant F(x_1)$. Por consiguiente, $F(x)$ es una función no decreciente.

2) $F(-\infty) = 0$, ya que el resultado del experimento $z \in (-\infty, x]$ se convierte en suceso imposible para $x \to -\infty$.

3) $F(+\infty) = 1$, ya que el resultado del experimento $z \in (-\infty, x]$ se convierte en suceso cierto para $x \to +\infty$.

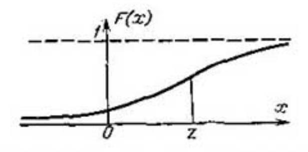


Fig. 5-2. Función de distribución de probabilidades para una magnitud aleatoria continua

En la figura 5-2 se presenta la vista general de la función que satisface las propiedades estipuladas.

c) Densidad de distribución de probabilidades

Conociendo la función de distribución de probabilidades puede calcularse la probabilidad de que el valor de la magnitud aleatoria se encuentre dentro del pequeño intervalo de x a $x + \Delta x$. Según (5-44) dicha probabilidad es igual a:

$$F(x + \Delta x) - F(x) = \frac{F(x + \Delta x) - F(x)}{\Delta x} \Delta x. \quad (5\text{-}45)$$

El primer factor del segundo miembro de esta expresión representa el valor de la probabilidad correspondiente a la unidad de longitud del sector Δx. Si existe el límite de esta relación con $\Delta x \to 0$, lo que se cumple, por lo común, para la mayoría de las magnitudes aleatorias continuas, el mismo se designa por $w(x)$ y se llama *densidad de distribución de las probabilidades*. Así pues,

$$w(x) = \lim_{\Delta x \to 0} \frac{F(x + \Delta x) - F(x)}{\Delta x} = \frac{dF(x)}{dx}, \quad (5\text{-}46)$$

de modo que

$$F(x + \Delta x) - F(x) = w(x)\,\Delta x. \quad (5\text{-}47)$$

En la figura 5-3 se muestra el carácter general de la gráfica de la función $w(x)$.

Aclaremos algunas propiedades de la función $w(x)$. Sean a y b puntos arbitrarios del eje real, siendo $b > a$. Hallemos el signi-

ficado de la integral $\int_a^b w(x)\,dx$ Sustituyendo de manera puramente formal $w(x)dx$ por $dF(x)$, obtenemos:

$$\int_a^b w(x)\,dx = \int_{F(a)}^{F(b)} dF(x) = F(b) - F(a). \tag{5-48}$$

Pero según (5-44) esta expresión da la probabilidad de que la magnitud aleatoria z tome un valor que esté dentro del intervalo (a, b). Así

$$P(a < z < b) = \int_a^b w(x)\,dx. \tag{5-49}$$

De esta fórmula se desprende como caso particular la correlación

$$\int_{-\infty}^{+\infty} w(x)\,dx = 1, \tag{5-50}$$

la cual muestra que la superficie limitada por la curva de la densidad de distribución de probabilidades $w(x)$ y el eje de abscisas es siempre igual a la unidad.

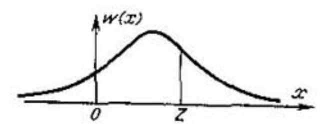


Fig. 5-3. Gráfica de densidades de la distribución de probabilidades

La fórmula (5-49) permite expresar de modo integral la relación entre las funciones $F(x)$ y $w(x)$. Comparando (5-44) con (5-49), hallamos que

$$F(x) = \int_{-\infty}^{x} w(\xi)\,d\xi. \tag{5-51}$$

Los conceptos de función de distribución de probabilidades y densidad de distribución de probabilidades también se extienden fácilmente al caso de una magnitud aleatoria unidimensional. Sin detenernos en esta cuestión, ocupémonos del análisis de algunas leyes de la distribución de magnitudes aleatorias continuas.

d) Distribución uniforme

Si una magnitud aleatoria con igual probabilidad puede tomar valores cualesquiera dentro de los límites del sector del eje real de α a β y no los puede tomar fuera de dicho sector, es decir,

tiene una densidad de distribución de las probabilidades

$$w(z)=\begin{cases}\dfrac{1}{\beta-\alpha}, & \alpha\leqslant z\leqslant\beta;\\ 0, & z<\alpha,\quad z>\beta,\end{cases} \tag{5-52}$$

la misma se llama distribuida uniformemente en el intervalo $[\alpha, \beta]$.

Es cómodo caracterizar la distribución uniforme por los parámetros $\nu=(\alpha+\beta)/2$ y $\omega=\beta-\alpha$ denominados *valor medio* y *amplitud* de la distribución y designar la distribución uniforme por $R(\nu, \omega)$. La densidad de distribución de las probabilidades $w(z)$ se expresa fácilmente mediante los parámetros de distribución $R(\nu, \omega)$ en la forma

$$w(z)=\begin{cases}\dfrac{1}{\omega}, & \nu-\dfrac{\omega}{2}\leqslant z\leqslant\nu+\dfrac{\omega}{2};\\ 0, & z<\nu-\dfrac{\omega}{2},\quad z>\nu+\dfrac{\omega}{2}.\end{cases} \tag{5-53}$$

Ejemplo 5-5. Al redondear números hasta valores enteros, los errores de redondeo pueden tomar con igual probabilidad valores de $-0{,}5$ a $+0{,}5$, es decir, obedecen a la distribución $R(0, 1)$.

Ejemplo 5-6. Al tomar lectura en una escala con valor de una división de δ los errores de lectura serán magnitudes aleatorias con distribución $R(0, \delta)$.

Ejemplo 5-7. Si un trolebús circula con intervalo de 10 min, para el pasajero que no conoce el horario, el tiempo de espera será una magnitud aleatoria con distribución $R(5, 10)$.

e) Distribución normal

La ley más importante de distribución de una magnitud aleatoria es la ley de distribución normal o de Gauss. Esta es la que se encuentra más a menudo en la práctica. Además, la ley de distribución normal es la ley límite a la cual se aproxima una serie de otras leyes de distribución al existir condiciones típicas que se presentan con gran frecuencia.

La ley normal se caracteriza por una función de distribución de las probabilidades del tipo

$$w(z)=\frac{1}{\sqrt{2\pi\sigma^2}}\,e^{-\frac{(z-\nu)^2}{2\sigma^2}} \tag{5-54}$$

para $-\infty<z<+\infty$. La distribución normal se determina por los parámetros ν, σ^2 denominados *media* y *dispersión* que en lo sucesivo van a designarse por $N(\nu, \sigma^2)$. En la figura 5-4 se muestra la gráfica de la función (5-54) para $\nu=1{,}0$ y algunos valores de σ.

La determinación de la probabilidad de que una magnitud aleatoria se encuentre en el intervalo dado (a, b) del eje real constituye una importantísima tarea de la teoría de las probabilidades.

En el caso de la distribución normal esta probabilidad es igual a:

$$P(a < z < b) = \frac{1}{\sqrt{2\pi}\,\sigma^2} \int_a^b e^{-\frac{(z-\nu)^2}{2\sigma^2}}\,dz. \qquad (5\text{-}55)$$

Esta expresión adquiere una forma más adecuada si se realiza un cambio de variable designándola por $(z-\nu)/\sigma = t$.

Con esto $dt = dz/\sigma$ y

$$P(a < z < b) = \frac{1}{\sqrt{2\pi}} \int_{\frac{a-\nu}{\sigma}}^{\frac{b-\nu}{\sigma}} e^{-\frac{1}{2}t^2}\,dt. \qquad (5\text{-}56)$$

Pero la integral del tipo $\int e^{-\frac{1}{2}t^2}dt$ no se expresa mediante funciones elementales. Por eso para el cálculo de (5-56) se emplean tablas de la función especial:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{1}{2}t^2}\,dt, \qquad (5\text{-}57)$$

denominada *integral de probabilidades* que se ofrecen en la mayoría de los libros sobre la teoría de las probabilidades. Tomando en

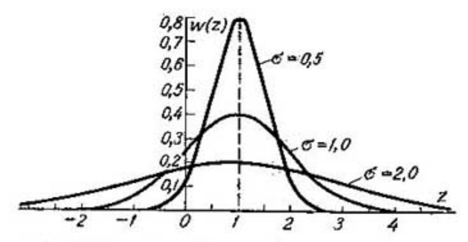


Fig. 5-4. Gráfica de $w(z)$ para la distribución normal

consideración (5-57) la probabilidad de que la magnitud aleatoria se halle en el intervalo (a, b) será igual a:

$$P(a < z < b) = \Phi\left(\frac{b-\nu}{\sigma}\right) - \Phi\left(\frac{a-\nu}{\sigma}\right). \qquad (5\text{-}58)$$

Al efectuar los cálculos por la fórmula (5-58) es conveniente recordar las siguientes propiedades de la integral de probabilidades:

$$\Phi(0) = 0; \quad \Phi(+\infty) = \frac{1}{2}; \quad \Phi(-x) = -\Phi(x). \qquad (5\text{-}59)$$

5-5. CARACTERÍSTICAS NUMERABLES DE LAS MAGNITUDES ALEATORIAS

a) Concepto de características numerables

La función de distribución o la densidad de distribución de probabilidades son las características más completas de las magnitudes aleatorias. Sin embargo, en muchos problemas prácticos resulta difícil o casi imposible representar la función de distribución de probabilidades.

Mientras tanto, para resolver diversos problemas es suficiente conocer sólo ciertos parámetros que caracterizan la magnitud aleatoria desde uno u otro punto de vista. Los parámetros numerables más difundidos que han recibido el nombre de características numerables o momentos numerables de las magnitudes aleatorias son el valor medio y la dispersión o desviación media cuadrática. Estos se determinan fácilmente con los datos experimentales y permiten discernir en rasgos generales sobre el carácter de la distribución de la magnitud aleatoria.

b) Valor medio (esperanza matemática) de una magnitud aleatoria

En muchos problemas la magnitud aleatoria considerada como resultado posible de cierto experimento se expresa por medio de un número real. Como ejemplos de las magnitudes aleatorias, que toman un valor numérico, pueden servir:

la cantidad de la corriente consumida por los moradores de una casa;

el valor de la tensión de la red en un momento dado;

la temperatura del aire en un momento dado del día;

el número de partículas cósmicas que llegan a la superficie terrestre en la unidad de tiempo;

el número de tantos obtenidos en el tiro al blanco por un tirador dado;

el número de compradores en una tienda en un momento dado, etc.

Para todos estos casos, su valor medio o esperanza matemática es una importante característica numérica de la magnitud aleatoria.

Supongamos que los elementos del espacio de resultados del experimento $Z = \{z_1, \ldots, z_m\}$ admiten estimación numérica. Si con N pruebas el resultado Z_1 sucedió r_1 veces, ..., el resultado z_m ocurrió r_m veces, de suerte que $r_1 + \ldots + r_m = N$, el valor medio de la magnitud aleatoria z que se designa por $M(z)$ o $\bar{z}$, es igual a:

$$M(z) = \bar{z} = \frac{z_1 r_1 + \ldots + z_m r_m}{N} = \sum_{i=1}^{m} z_i \frac{r_i}{N}, \qquad (5\text{-}60)$$

Si el número de pruebas N fue suficientemente grande, la relación $r_i/N = p_i$ puede considerarse como la probabilidad del resultado z_i. Entonces, omitiendo los subíndices, la fórmula para determinar el valor medio puede escribirse en la forma

$$M(z) = \bar{z} = \sum_{z \in Z} zp(z). \tag{5-61}$$

El valor medio $M(z)$ determinado por la fórmula (5-61) se denomina comúnmente *esperanza matemática* de la magnitud aleatoria z.

La fórmula (5-61) puede servir para determinar el valor medio en caso de que el espacio de resultados de la prueba Z sea un conjunto finito. Ahora bien, si z es una magnitud aleatoria continua con densidad de distribución de probabilidades $w(z)$, por $p(z)$ en la fórmula (5-61) puede entenderse la probabilidad de que el valor de la magnitud aleatoria se halle entre los límites de z a $z + \Delta z$, es decir, suponerse que $p(z) = w(z)\Delta z$. Pasando al límite con $\Delta z \to 0$ y sustituyendo, respectivamente, Δz por dz y la suma por la integral, obtenemos:

$$M(z) = \bar{z} = \int_{-\infty}^{+\infty} zw(z)\,dz. \tag{5-62}$$

c) Valor medio de una función de magnitud aleatoria

En la práctica, a menudo encontramos problemas en los que la magnitud aleatoria no se expresa con un número. Así pues, la producción elaborada por una fábrica se divide en útil y defectuosa. Al lanzar una moneda, el resultado del experimento puede ser "cara" o "cruz". Una carta sacada de la baraja al azar se caracteriza por su nombre y palo. La señal radar recibida contiene información sobre la presencia o ausencia del objetivo, etc. En todos estos ejemplos las magnitudes aleatorias presentan diferencia no cuantitativa sino sólo cualitativa. No obstante, por lo común es posible introducir también la estimación cuantitativa para las magnitudes aleatorias que solamente presentan diferencias cualitativas.

Consideramos la magnitud aleatoria como el resultado de cierto experimento que se realiza artificialmente o es el resultado de un proceso natural. Con esto hay que recordar que cualquier experimento requiere algunos gastos y se realiza no para satisfacer nuestra curiosidad, sino para alcanzar un objetivo determinado. Uno u otro resultado del experimento es deseable o indeseable dependiendo de hasta qué punto se logra el objetivo propuesto. La consecución del objetivo deseado puede considerarse como ganancia o ventaja y su no consecución, como pérdida o desventaja. Tanto lo ganado como lo perdido puede expresarse con números que representan, por ejemplo, ciertas sumas de dinero en rublos.

De esta manera, a cada resultado de la experiencia $z \in Z$ puede ponerse en correspondencia cierta estimación numérica, o sea, puede efectuarse el reflejo f del espacio de resultados del experimento Z en el conjunto de números reales R

$$f : Z \to R. \tag{5-63}$$

Este reflejo proporciona la función real $f(z)$ definida en Z, con la que podemos operar como con una magnitud aleatoria y, en particular, determinar su valor medio.

Sea $f(z)$ cierta función numérica definida en el conjunto Z. Esto significa que a los elementos $z_1, \ldots, z_m$ del conjunto Z corresponden los valores $f(z_1), \ldots, f(z_m)$ de la función. Las probabilidades de estos valores serán las mismas que las de las magnitudes aleatorias $z_1, \ldots, z_m$. De este modo, $p(z)$ representa la distribución de las probabilidades tanto para la magnitud aleatoria z como, asimismo, para la función de la magnitud aleatoria $f(z)$. Por consiguiente, el valor medio de la función de una magnitud aleatoria puede encontrarse por las fórmulas para el valor medio de una magnitud aleatoria gracias a la sustitución de z por $f(z)$, es decir,

$$M[f(z)] = \overline{f(z)} = \sum_{z \in Z} f(z)\, p(z) \tag{5-64}$$

para una magnitud aleatoria discreta, y

$$M[f(z)] = \overline{f(z)} = \int_{-\infty}^{+\infty} f(z)\, w(z)\, dz \tag{5-65}$$

para una continua.

Notemos que la magnitud $\overline{f(z)}$ calculada por la fórmula (5-64) depende de la distribución de probabilidades $p(z)$. Esto quiere decir que si cambia la distribución de probabilidades $p(z)$ en el espacio Z, cambia también el valor medio de la función $\overline{f(z)}$. Para recalcar esta circunstancia vamos a designar a veces por $f(p)$ el valor medio de la función $\overline{f(z)}$, suponiendo que

$$f(p) = \overline{f(z)}. \tag{5-66}$$

El empleo en el caso dado, para designar el valor medio de la función, de la misma letra f que para denotar la propia función, es admisible y no puede provocar confusiones ya que las funciones $f(z)$ y $f(p)$ están definidas en conjuntos diferentes: la función $f(z)$ en el espacio de resultados del experimento Z, y $f(p)$ en el espacio de todas las distribuciones posibles de probabilidades $p(z)$.

Ejemplo 5-8. La probabilidad de que el aparato producido por una fábrica resulte defectuoso es igual a p.

¿Cuál es la ganancia media correspondiente a un aparato, si a es el costo de producción y b, el precio del mismo?

Utilizando la fórmula (5-64), hallamos que:

$$f(p) = (b - a)(1 - p) - ap = b(1 - p) - a.$$

En muchos problemas el valor medio se considera como la ganancia media con un gran número de experimentos. En este caso se considera racional llevar a cabo el experimento si $\overline{f(z)} = f(p) \geqslant 0$.

Otra interpretación del valor medio se refiere a las apuestas. Supongamos que cierto suceso puede ocurrir con la probabilidad p. El hacer una apuesta presupone que uno de los participantes está de acuerdo en pagar la suma b si no occure un suceso a condición de que el otro participante le pague la suma a si éste ocurre. De este modo, el primer participante recibe la suma a con la probabilidad p y paga la suma b con la probabilidad $1 - p$. Su ganancia media es

$$f(p) = ap - b(1 - p). \tag{5-67}$$

La apuesta se considera equitativa si la ganancia media resulta igual a cero, o sea, bajo la condición de que

$$\frac{a}{b} = \frac{1 - p}{p}. \tag{5-68}$$

Teorema 5-1 (teorema del valor medio). *Utilizando para determinar $f(p)$ la expresión (5-64) y sustituyendo en cada sumando $f(z)$ por $\min_z f(z)$, obtenemos la correlación*

$$f(p) \geqslant \min_z f(z) \sum_{z \in Z} p(z) = \min_z f(z), \tag{5-69}$$

que viene a expresar el teorema del valor medio.

d) Valor medio de una función de dos magnitudes aleatorias

En muchos casos se requiere operar no con uno sino con varios sucesos aleatorios e introducir una función real que depende de los resultados de cada uno de estos sucesos. Así, por ejemplo, en las competiciones de atletismo el resultado de cada participante es un suceso aleatorio. Ahora bien, el lugar ocupado por el equipo es función tanto de los resultados de los miembros de su propio equipo, como de los resultados de los miembros del equipo rival.

Nos limitaremos al análisis de dos magnitudes aleatorias independientes para las que designaremos, respectivamente, por X e Y los espacios de resultados del experimento. Designemos por $p(x)$ y $q(y)$ la distribución de las probabilidades en los espacios X e Y.

Sea $f(x, y)$ una función real determinada para $x \in X$ e $y \in Y$ cualesquiera. Entonces se dice que la función $f(x, y)$ está presentada en el producto directo de los conjuntos $X \times Y$. El valor medio de esta función dependerá del carácter de las distribuciones de probabilidades $p(x)$ y $q(y)$ así que puede escribirse:

$$M[f(x, y)] = \overline{f(x, y)} = f(p, q). \tag{5-70}$$

Para determinar $f(p, q)$ presentaremos cierta $y \in Y$ determinada. Con la y propuesta la función $f(x, y)$ solamente lo será de x

y su valor medio, que sólo dependerá de $p(x)$, puede designarse por $f_p(y)$ y se halla por la fórmula (5-64):

$$M_x[f(x, y)] = f_p(y) = \sum_x f(x, y)\, p(x). \qquad (5\text{-}71)$$

La magnitud $f_p(y)$ ya no depende de x sino solamente de y. La promediación por y de esta magnitud da $f(p, q)$:

$$f(p, q) = M_y[f_p(y)] = \sum_y f_p(y)\, q(y) = \sum_{x,\, y} f(x, y)\, p(x)\, q(y). \qquad (5\text{-}72)$$

e) Esperanza matemática condicional

Sea Z el espacio de resultados del experimento y $f(z)$, cierta función numérica presentada en el conjunto Z cuyo valor medio o esperanza matemática se determina por la expresión (5-64).

Supongamos ahora que ocurrió el suceso $S \subseteq Z$. Entonces la distribución de probabilidades $p(z)$ se sustituye por la distribución condicional de probabilidades $p(z|S)$. Si en la fórmula para el valor medio ponemos en lugar de $p(z)$ la magnitud $p(z|S)$, obtenemos la esperanza matemática condicional de la función $f(z)$:

$$M(f|S) = \sum_{z \in Z} f(z)\, p(z|S) = \frac{1}{P(S)} \sum_{z \in Z} f(z)\, p(z \cap S) = \frac{\sum_{z \in S} f(z)\, p(z)}{\sum_{z \in S} p(z)}. \qquad (5\text{-}73)$$

f) Propiedades del valor medio

1. El valor medio de una magnitud no aleatoria es igual a la misma magnitud.

Sea c una magnitud no aleatoria, es decir, constante. Entonces c puede considerarse como el resultado de un experimento en el cual el espacio de resultados Z consta de un solo elemento c de suerte que $p(c) = 1$. Por la fórmula (5-61) hallamos que

$$M(c) = cp(c) = c. \qquad (5\text{-}74)$$

2. El factor constante puede sacarse del signo de valor medio

$$M(cz) = \sum_{z \in Z} czp(z) = c \sum_{z \in Z} zp(z) = cM(z). \qquad (5\text{-}75)$$

3. El valor medio de una suma de magnitudes aleatorias es igual a la suma de sus valores medios.

Sean x e y magnitudes aleatorias que tienen espacios de resultados del experimento X e Y y distribuciones de probabilidades en estos espacios $p(x)$ y $q(y)$. Designemos $f(x, y) = x + y$. Por la

fórmula (5-72) encontramos que

$$\overline{x+y}=\sum_{x,y}(x+y)p(x)q(y)=\sum_{x}xp(x)\sum_{y}q(y)+$$
$$+\sum_{y}yq(y)\sum_{x}p(x)=\sum_{x}xp(x)+\sum_{y}yq(y)=\bar{x}+\bar{y}. \qquad (5\text{-}76)$$

4. La esperanza matemática de una magnitud aleatoria centrada es igual a cero.

Sea x una magnitud aleatoria y ν, su valor medio. La magnitud aleatoria $\overset{\circ}{z}=z-\nu$ se denomina magnitud aleatoria centrada. Para la magnitud aleatoria centrada tenemos:

$$M(\overset{\circ}{z})=M(z-\nu)=M(z)-\nu=0. \qquad (5\text{-}77)$$

g) Momentos. Dispersión. Desviación cuadrática media

Más importantes características numéricas de una magnitud aleatoria son sus momentos que se dividen en iniciales y centrales.

El momento inicial de orden s de la magnitud aleatoria z se determina por las fórmulas:

$$\alpha_s(z)=\sum_{z}z^s p(z) \qquad (5\text{-}78)$$

para una magnitud aleatoria discreta, y

$$\alpha_s(z)=\int_{-\infty}^{+\infty}z^s w(z)\,dz \qquad (5\text{-}79)$$

para una magnitud aleatoria continua.

No es difícil observar que el valor medio $M(z)=\bar{z}$, designado a continuación por ν o ν_z, constituye el momento inicial de primer orden $\alpha_1(z)$. El momento inicial de segundo orden es el valor medio del cuadrado de la magnitud aleatoria

$$\alpha_2(z)=\sum_{z}z^2 p(z)=M(z^2)=\overline{z^2}. \qquad (5\text{-}80)$$

Si en calidad de magnitud aleatoria se utiliza la magnitud aleatoria centrada $\overset{\circ}{z}=z-\nu$, las fórmulas (5-78) y (5-79) brindan expresiones para los momentos centrales de orden s, designados por $\mu_s(z)$. Como se ve en la correlación (5-77) el momento central de primer orden es igual a cero.

Sirve como importante característica numérica de una magnitud aleatoria el momento central de segundo orden llamado *dispersión* de la magnitud aleatoria z que se designa por $D(z)$ o D_z. Las fórmulas para determinar la dispersión tienen el aspecto:

$$D(z)=\mu_2(z)=\sum_{z}(z-\nu_z)^2 p(z) \qquad (5\text{-}81)$$

para una magnitud aleatoria discreta y

$$D(z)=\mu_2(z)=\int_{-\infty}^{+\infty}(z-\nu_z)^2\,w(z)\,dz \qquad (5\text{-}82)$$

para una magnitud aleatoria continua. La dispersión caracteriza la desviación de los valores aislados de la magnitud aleatoria respecto al valor medio, o sea, constituye la característica de dispersión de una magnitud aleatoria. Cuanto menor es la dispersión, tanto más estrechamente se concentran en las cercanías del valor medio los valores aislados de la magnitud aleatoria.

En una serie de casos la dispersión no resulta cómoda para su empleo práctico ya que tiene una dimensionalidad igual al cuadrado de la magnitud aleatoria. Por eso, se utiliza frecuentemente como característica de dispersión de una magnitud aleatoria, la raíz cuadrada de la dispersión que ha recibido el nombre de *desviación cuadrática media* que se designa por σ o σ_z:

$$\sigma_z=\sqrt{D(z)}. \qquad (5\text{-}83)$$

Anotemos algunas propiedades de la dispersión.

1. La dispersión de una magnitud no aleatoria es igual a cero:

$$D(c)=M[(c-\bar{c})^2]=M(0)=0. \qquad (5\text{-}84)$$

2. La magnitud no aleatoria puede sacarse del signo de depresión elevándola al cuadrado:

$$D(cz)=M[(cz-c\nu)^2]=c^2D(z). \qquad (5\text{-}85)$$

Más adelante estudiaremos algunas otras propiedades de la dispersión.

h) Regresión y correlación

En los casos en que es necesario trabajar con varias magnitudes aleatorias, por ejemplo, con dos, el valor medio y la dispersión pueden caracterizar por separado cada una de estas magnitudes. No obstante, en semejantes casos es muy importante conocer la influencia de una magnitud sobre la otra, o sea, tomar en consideración el carácter de la conexión entre las magnitudes aleatorias. La regresión y la correlación sirven precisamente para expresar esta conexión de manera cuantitativa.

Sean x e y dos magnitudes aleatorias similares a aquellas que se estudiaron en el § 5-3. Para concretar, vamos a suponer que x representa la altura, e y, el diámetro de los árboles en una parcela de bosque. A cada árbol va a corresponder un punto en el plano (x, y), y la totalidad de los árboles se representará por el conjunto de los puntos mostrados en la figura 5-5. En el caso examinado las magnitudes $\nu_x=\bar{x}$ y $\nu_y=\bar{y}$ dan el valor medio de la altura

y del diámetro de los árboles, y σ_x, σ_y caracterizan la dispersión de la altura y del diámetro con respecto a los valores medios.

Junto con el análisis de la altura y del diámetro medios, presenta gran interés el estudio de la variación del diámetro en dependencia de la altura del árbol. Sin embargo, el diámetro y es una magnitud aleatoria para los árboles que tienen la misma altura x. Por eso sólo puede hablarse sobre la dependencia del valor medio del diámetro $\bar{y}$ respecto a la altura x, o sea, determinarse la magnitud $\bar{y}(x)$ que constituye el valor medio condicional $M(y|x)$. Empleando (5-73) y designando por $p(x, y)$ la probabilidad compatible de los valores dados x e y, hallamos:

$$\bar{y}(x) = M(y|x) = \frac{\sum_y yp(x, y)}{\sum_y p(x, y)}. \quad (5\text{-}86)$$

Fig. 5-5. Línea de regresión $\bar{y}(x)$

Determinando $\bar{y}(x)$ para diversas x, podemos construir, como se muestra en la figura 5-5, la línea que expone gráficamente esta dependencia y que se denomina *línea de regresión* de y por x. Análogamente puede obtenerse la dependencia $\bar{x}(y) = M(x|y)$ llamada regresión de x por y.

Nos limitaremos a estudiar la regresión para el caso de la regresión lineal, el más sencillo y que se encuentra con la mayor frecuencia en la práctica, es decir, cuando la línea de regresión es una línea recta cuya ecuación puede escribirse en la forma

$$\bar{y}(x) = a + b(x - \bar{x}). \quad (5\text{-}87)$$

Elijamos los coeficientes a y b de tal modo que obtengamos la mayor concentración de puntos (x, y) en las cercanías de la recta $\bar{y}(x)$ lo que puede expresarse por la condición

$$\psi(a, b) = M\{[y - \bar{y}(x)]^2\} = \text{mín}. \quad (5\text{-}88)$$

Teniendo en cuenta (5-87) la condición (5-88) brinda el siguiente sistema de ecuaciones para determinar a y b:

$$\left.\begin{aligned} \bar{y} - a = 0; \\ M[y(x - \bar{x})] - b\sigma_x^2 = 0. \end{aligned}\right\} \quad (5\text{-}89)$$

Denominemos covarianza entre x e y la magnitud

$$\mu_{xy} = \text{cov}(x, y) = M[(y - \bar{y})(\bar{x} - \bar{x})], \quad (5\text{-}90)$$

Es fácil ver que μ_{xy} puede representarse como

$$\mu_{xy} = M[y(x - \bar{x}) - \bar{y}(x - \bar{x})] =$$
$$= M[y(x - \bar{x})] - \bar{y}M(x - \bar{x}) = M[y(x - \bar{x})]. \quad (5\text{-}91)$$

Tomando en consideración (5-91), de (5-89) encontramos los valores a y b que definen las líneas de regresión:

$$a = \bar{y}; \quad b = \frac{\mu_{xy}}{\sigma_x^2}. \tag{5-92}$$

La covarianza μ_{xy} puede servir de medida de la correlación entre las magnitudes aleatorias x e y. Pero, por cuanto nos hemos limitado al análisis del caso en que la línea de regresión es una recta, la covarianza no va a caracterizar la correlación cualquiera, sino solamente la lineal, es decir, la tendencia de una magnitud aleatoria a crecer o decrecer en promedio según la ley lineal al crecer o decrecer otra magnitud aleatoria. Para representar con más claridad cómo depende la covarianza del carácter de la correlación entre las magnitudes aleatorias será útil examinar los casos extremos.

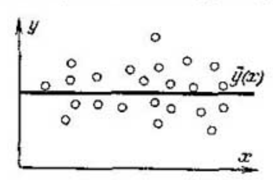


Fig. 5-6. Línea de regresión para las magnitudes aleatorias independientes

1. $\mu_{xy} = 0$. Esto significa que $b = 0$, $\bar{y} = a$ y la línea de regresión es horizontal como se muestra en la figura 5-6. Como vemos, en los puntos experimentales los valores y se agrupan alrededor del valor a independientemente del valor x. Luego, en el caso dado x e y son las magnitudes aleatorias independientes.

2. La correlación entre x e y será la más estrecha si dichas magnitudes están ligadas mediante una dependencia funcional lineal. Entonces, todos los puntos experimentales van a estar en la línea de regresión de tal modo que $y = \bar{y}(x)$. Sustituyendo $\bar{y}(x)$ por y transcribamos la ecuación de la línea de regresión en la forma

$$y - \bar{y} = \frac{\mu_{xy}}{\sigma_x^2}(x - \bar{x}). \tag{5-93}$$

Hallamos la dispersión σ_y^2 para este caso:

$$\sigma_y^2 = M[(y - \bar{y})^2] = \frac{\mu_{xy}^2}{\sigma_x^2}, \tag{5-94}$$

de donde

$$\mu_{xy} = \sigma_x \sigma_y. \tag{5-95}$$

Así pues, dependiendo del grado de correlación entre las magnitudes x e y, la covarianza μ_{xy} puede cambiar de 0 a $\sigma_x \sigma_y$.

Es un poco incómodo estimar con ayuda de la covarianza μ_{xy} cuán estrecho es el nexo entre las magnitudes aleatorias x e y, ya que μ_{xy} depende de la dispersión de éstas. Más cómodo resulta

utilizar con dicho objetivo el *coeficiente de correlación*

$$r_{xy}=\frac{\mu_{xy}}{\sigma_x\sigma_y}, \qquad (5\text{-}96)$$

que puede fluctuar entre los límites desde cero para las magnitudes aleatorias independientes, hasta la unidad, si las magnitudes aleatorias están relacionadas mediante una dependencia funcional lineal.

El empleo del concepto de covarianza permite obtener correlaciones sencillas para el valor medio del producto y la dispersión de la suma de dos magnitudes aleatorias.

Proponemos al lector comprobar por sí mismo que en estos casos se emplean las correlaciones

$$\overline{xy}=\bar{x}\bar{y}+\mu_{xy}; \qquad (5\text{-}97)$$

$$D(x+y)=D(x)+D(y)+2\mu_{xy}. \qquad (5\text{-}98)$$

Cuando las magnitudes aleatorias x e y son independientes,

$$\overline{xy}=\bar{x}\bar{y}; \qquad (5\text{-}99)$$

$$D(x+y)=D(x)+D(y). \qquad (5\text{-}100)$$

5-6. PROCESOS ESTOCASTICOS DISCRETOS

a) Tipos de procesos estocásticos discretos

Supongamos que cierto sistema que se halla bajo la influencia del total de condiciones exteriores puede alterar casualmente su estado en los momentos de tiempo discretos designados convencionalmente por 0, 1, ...; los intervalos entre ellos se denominan pasos. Entonces, a resultas del n-esimo paso el sistema toma el estado x_n, que podemos considerar como el resultado de cierto experimento y, por consiguiente, será elemento del conjunto de los resultados admisibles para este paso:

$$Z_n=\{z_0^n,\ z_1^n,\ \ldots,\ z_{L_n}^n\}.$$

Consideremos que el conjunto Z_0 consta de un solo elemento z_0^0 que representa el estado inicial del sistema. La secuencia de estados asumidos por el sistema x_0, x_1 ... se denomina proceso estocástico.

El proceso estocástico estará presentado si en cada paso está dada la distribución de probabilidades en el espacio de resultados que determina los estados a los que puede pasar el sistema como resultado del paso siguiente. Esta distribución de probabilidades depende, en el caso general, de los estados asumidos en todos los pasos anteriores, y para cualquier n, se determina por el juego de

probabilidades condicionales:

$$P(x_{n+1}=z_k^{n+1}\,|\,x_0, \ldots, x_n), \quad k=0, 1, \ldots, L_n, \qquad (5\text{-}101)$$

siendo

$$\sum_{k=0}^{L_n} p(x_{n+1}=z_k^{n+1}\,|\,x_0, \ldots, x_n)=1. \qquad (5\text{-}102)$$

El proceso estocástico va a continuar infinitamente si no lo rompemos artificialmente en algún paso. En la práctica siempre nos es necesario interrumpir el proceso después de un número finito determinado de pasos, es decir, operar con procesos *finitos* o *de pasos múltiples*. No obstante, los procesos infinitos son aproximaciones útiles de los procesos finitos que son más reales, máxime que a menudo éstos resultan más sencillos desde el punto de vista analítico.

Se denomina *no controlado* el proceso estocástico en el cual no podemos intervenir, o sea, influir en la distribución de probabilidades (5-101).

Según la índole de la distribución de probabilidades (5-101), los procesos estocásticos se dividen en una serie de tipos de los cuales los más difundidos son los procesos con valores independientes, los de pruebas independientes y las cadenas de Márkov.

Un proceso estocástico finito se denomina *proceso con valores independientes* si la distribución de probabilidades en el espacio de resultados durante cada paso no depende de los resultados de los pasos anteriores:

$$P(x_n=z_k^n\,|\,x_0, \ldots, x_{n-1})=p(x_n=z_k^n). \qquad (5\text{-}103)$$

Es cómodo designar esta distribución de probabilidades simplemente por $p_n(z)$, donde $z \in Z_n$, $n=0, 1 \ldots$

Una importante propiedad del proceso con valores independientes consiste en que cualquier secuencia de estados $x_0, x_1, \ldots, x_n$ puede considerarse como el total de los resultados independientes del experimento. Luego, la probabilidad de que aparezca esta secuencia será igual al producto de las probabilidades de los resultados que forman parte de ella, lo que puede escribirse en la forma

$$p(x_0, x_1, \ldots, x_n)=p_0(x_0)\,p_1(x_1)\ldots p_n(x_n),$$
$$x_i \in Z_i, \quad i=0, 1, \ldots, n. \qquad (5\text{-}104)$$

Un importante caso particular de procesos con valores independientes son los procesos de pruebas independientes. El *proceso de pruebas independientes* es un proceso con valores independientes en el cual el espacio de resultados del experimento $Z=\{z_0, z_1 \ldots, z_L\}$ y la distribución de probabilidades en este espacio son iguales en cada paso, es decir,

$$p_n(z)=p_m(z)=p(z) \qquad (5\text{-}105)$$

con cualquier $z \in Z$ y cualesquiera m y n. En otras palabras, el proceso de pruebas independientes constituye la múltiple repetición de un mismo experimento en las mismas condiciones.

Como en el caso de los procesos con valores independientes, la probabilidad de obtener una secuencia concreta de resultados del experimento $x_0, x_1, \ldots, x_n$ en el proceso de pruebas independientes, es igual al producto de las probabilidades de los resultados separados.

Un proceso estocástico se denomina *cadena de Márkov* si el espacio de estados $Z = \{z_0, z_1, \ldots, z_L\}$ es el mismo para cada paso y la distribución de probabilidades de los resultados en cada paso sólo depende de los resultados del paso anterior, de suerte que

$$p(x_{n+1} = z \mid x_0, \ldots, x_n) = p(x_{n+1} = z \mid x_n). \qquad (5\text{-}106)$$

Es conveniente introducir otras designaciones para las probabilidades del tipo (5-106). Supongamos que en el n-ésimo paso el sistema se hallaba en el estado $z_i \in Z$. Entonces la distribución de probabilidades (5-106) determinara la probabiladad de que en el paso $(n+1)$ el sistema se halle en el estado $z_j \in Z$. En el caso general dicha probabilidad depende del número del paso y puede designarse por $p_{ij}(n)$. De esta manera, las probabilidades

$$p_{ij}(n) = p(x_{n+1} = z_j \mid x_n = z_i), \qquad (5\text{-}107)$$

que satisface la condición

$$p_{ij}(n) \geqslant 0, \qquad \sum_{j=0}^{L} p_{ij}(n) = 1, \qquad (5\text{-}108)$$

son las probabilidades de transición del estado z_i al estado z_j en el paso n.

La cadena de Márkov se llama *no homogénea* en caso de que las probábilidades de transición $p_{ij}(n)$ dependan del número del paso n. Sin embargo, juegan un gran papel en diversas aplicaciones las cadenas de Márkov *homogéneas* en las que las posibilidades de transición se mantienen iguales para cualquier paso y pueden designarse simplemente por p_{ij}.

b) Proceso de pruebas independientes con dos resultados. Distribución binomial de probabilidades

Examinemos más detalladamente el caso en que en el proceso de pruebas independientes sólo nos interesan dos resultados de cada experimento, a saber, ocurrirá o no cierto suceso S como resultado de dicho experimento. Vamos a denominar éxito a un resultado del experimento y fracaso, al otro.

Designemos por p la probabilidad de éxito en un experimento aislado, y por $q = 1 - p$, la probabilidad de fracaso. Hagámonos la pregunta cuál es la probabilidad $w(r, n, p)$ de que con n experimentos vayan a ocurrir r éxitos.

Para mayor comodidad, designemos por 1 el éxito del experimento y por O, su fracaso. Con esto, al logro de r éxitos con n experimentos va a corresponder una secuencia de n resultados que consta de r unidades y $n - 1$ ceros. Como que los resultados de los éxitos separados son independientes, la probabilidad de esta secuencia es igual al producto de las probabilidades de dichos resultados separados y se determina por la expresión

$$p^r q^{n-r}. \tag{5-109}$$

El número total de distintas secuencias de n elementos que contienen r unidades, es igual al número de combinaciones de n elementos por r, es decir,

$$\binom{n}{r} = \frac{n!}{r!\,(n-r)!}. \tag{5-110}$$

Ya que todas estas secuencias son incompatibles (si apareció una de ellas, queda excluida la posibilidad de que surja cualquier

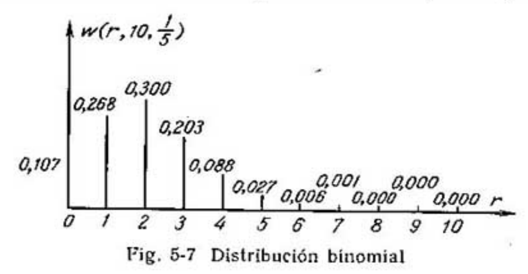


Fig. 5-7 Distribución binomial

otra), la probabilidad de que aparezca cualquiera de ellas será igual a:

$$w(r,\ n,\ p) = \binom{n}{r} p^r q^{n-r}. \tag{5-111}$$

Es fácil comprobar que con esto la probabilidad suma de que aparezca cualquier número de sucesos de 0 a n es igual a la unidad.

La distribución de probabilidades determinada por la fórmula (5-111) se llama *distribución binomial*. Para representar claramente la naturaleza de la distribución binomial, en la figura 5-7 se muestra la gráfica de dependencia de las ordenadas de esta distribución respecto a r para $n = 10$ y $p = 1/5$, es decir, la gráfica de distribución $w(r,\ 10,\ 1/5)$.

Muy a menudo, es necesario determinar el valor r con el que la distribución binomial alcanza el máximo. Calculemos previa-

mente la relación de dos ordenadas adyacentes de la distribución binomial

$$\frac{w(r, n, p)}{w(r-1, n, p)} = \frac{\binom{n}{r} p^r q^{n-r}}{\binom{n}{r-1} p^{r-1} q^{n-r+1}} = \frac{(n-r+1)p}{rq}. \qquad (5\text{-}112)$$

Esta relación no será menor de la unidad si $r \geqslant (n+1)p$. Luego, la función $w(r, n, p)$ alcanza su máximo dado

$$r = \text{ent}\,(n+1)\,p, \qquad (5\text{-}113)$$

donde se ha designado por ent x la parte entera del número x.

Sirven de características numéricas de la distribución binomial el valor medio ν_r y la dispersión σ_r^2 que se determinan por las correlaciones

$$\nu_r = np \qquad \sigma_r^2 = npq. \qquad (5\text{-}114)$$

c) Distribución de Poisson

Los valores $w(r, n, p)$ son de difícil cálculo en caso de ser grandes las n. Sin embargo, hay métodos aproximados que simplifican notablemente el cálculo de estas probabilidades. Uno de ellos es sustituir la distribución binomial por la de Poisson.

Sea $w(r, n, p)$ la distribución binomial. Es fácil cerciorarse de que el límite de esta distribución con $n \to \infty$ y $p \to 0$ a condición de que $np = a =$ const, de suerte que $p = a/n$ es,

$$\lim_{n\to\infty} w(r, n, n/a) = \frac{a^r e^{-a}}{r!}. \qquad (5\text{-}115)$$

La distribución del tipo (5-115) se designa por $w(r, a)$ y se llama *distribución de Poisson*. De esta manera,

$$w(r, a) = \frac{a^r e^{-a}}{r!}. \qquad (5\text{-}116)$$

Constituye una particularidad característica de la distribución de Poisson el hecho de que el valor medio y la dispersión de esta distribución son iguales y valen a a.

La distribución de Poisson constituye una buena aproximación a la distribución binomial con pequeñas p. Así, para $p \leqslant 0{,}01$ puede sustituirse el cálculo de la distribución binomial por el cómputo de la distribución de Poisson comenzando por $n = 10$.

Aunque hemos llegado a la distribución de Poisson como al caso extremo de distribución binomial, no obstante, ella representa una distribución de clase independiente de los procesos estocásticos, denominada clase de *fenómenos aleatorios raros*.

Examinemos la secuencia de sucesos S que siguen uno tras otro al cabo de intervalos de tiempo aleatorios. El número de estos

sucesos que ocurren en un intervalo de duración τ será una magnitud aleatoria con valor medio de $\lambda\tau$ donde λ es el número medio de sucesos en la unidad de tiempo. Impongámonos la tarea de determinar la probabilidad de que en el intervalo τ ocurran exactamente r sucesos.

Vamos a considerar raros los sucesos S en el sentido de que puede dividirse el intervalo τ en pequeños intervalos de duración $\Delta\tau$ en cada uno de los cuales no puede ocurrir más de un suceso S. El número total de estos intervalos es igual a $n = \tau/\Delta\tau$ y la probabilidad de que el suceso ocurra en uno de los intervalos $\Delta\tau$, igual a $p = \lambda\tau/n$, de modo que $np = \lambda\tau$. Por cuanto puede considerarse que en cada uno de los intervalos $\Delta\tau$ el suceso S ocurre independientemente de si acaeció o no en otros intervalos, arribamos al proceso de pruebas independientes con distribución de probabilidades $w(r, n, p)$, donde $np = \lambda\tau$.

Sin embargo, la suposición de que en el intervalo $\Delta\tau$ no puede ocurrir más de un suceso S, sólo quedará justificada en el caso en que la duración de dicho intervalo sea muy pequeña, es decir, en el límite con $\Delta\tau \to 0$. Entonces, como $n \to \infty$ y $p \to 0$, dado $np = \lambda\tau = \text{const}$, el proceso de los fenómenos raros obedece a la ley de Poisson con la distribución de probabilidades

$$w(r, \lambda\tau) = \frac{(\lambda\tau)^r e^{-\lambda\tau}}{r!}. \qquad (5\text{-}117)$$

d) Distribución exponencial. Concepto de fiabilidad

En muchos problemas es de interés la distribución de probabilidades para los intervalos de tiempo entre los sucesos aleatorios en el proceso de pruebas independientes. De la fórmula (5-117) para $r = 0$ y $r = 1$ obtenemos:

$e^{-\lambda\tau}$ es la probabilidad de que en el intervalo de duración τ no acontecerá ningún suceso;

$\lambda\tau e^{-\lambda\tau}$ es la probabilidad de que en el intervalo de duración τ ocurra exactamente un suceso.

La probabilidad de que en el intervalo de duración τ ocurra más de un suceso, es igual a:

$$1 - (e^{-\lambda\tau} + \lambda\tau e^{-\lambda\tau}) = 1 - \left\{\left[1 - \lambda\tau + \frac{(\lambda\tau)^2}{2!} - \dots\right] - \right.$$
$$\left. - \lambda\tau\left[1 - \lambda\tau + \frac{(\lambda\tau)^2}{2!} - \dots\right]\right\} = \frac{(\lambda\tau)^2}{2} + \dots \qquad (5\text{-}118)$$

Como vemos, esta probabilidad es proporcional a la magnitud $(\lambda\tau)^2$, lo que significa que con pequeñas $\lambda\tau$ podemos despreciar la probabilidad de que tengan lugar dos o más sucesos en el intervalo τ. Entonces, la probabilidad de que acontezca un suceso en el intervalo τ será igual a:

$$F(\lambda, \tau) = 1 - e^{-\lambda\tau}, \quad \tau \geqslant 0. \qquad (5\text{-}119)$$

Dicho con propiedad, la expresión (5-119) determina la probabilidad de que el tiempo en el cual ocurre un suceso no exceda de τ. Diferenciando (5-119) por τ obtendremos la densidad de distribución para esta probabilidad que representa la probabilidad de que el suceso ocurra en el momento τ:

$$w(\lambda, \tau) = \lambda e^{-\lambda\tau}, \quad \tau \geqslant 0. \tag{5-120}$$

Conociendo la densidad de distribución para τ es fácil determinar el tiempo medio durante el cual no ocurre ningún suceso:

$$\tau_{\text{med}} = \int_0^\infty \tau w(\lambda\tau)\, d\tau = \frac{1}{\lambda}. \tag{5-121}$$

Las fórmulas (5-119) y (5-120) definen la distribución *exponencial* que juega importante papel en la teoría de la fiabilidad.

Supongamos que cierto dispositivo consta de un gran número de elementos, cada uno de los cuales puede averiarse con el tiempo. La avería del elemento se llama su *falla*. En la mayoría de los casos las fallas de algunos elementos ocurren independientemente unas de otras y su secuencia puede considerarse como un proceso de sucesos independientes en el cual λ determina el número de fallas en la unidad de tiempo y se llama *intensidad* de fallas y la magnitud $1/\lambda$ según (5-121) determina el tiempo medio de funcionamiento sin fallas.

Habitualmente se emplea como característica de fiabilidad del equipo, la magnitud $h(\tau)$ denominada *peligro de falla*, que expresa la densidad de distribución de la probabilidad de falla en el momento τ bajo la condición de que no han ocurrido fallas hasta ese momento.

Definiremos previamente la magnitud $h(\tau)\Delta\tau$ que da la probabilidad de que la falla ocurrirá en el intervalo $\Delta\tau$ bajo la condición de que con anterioridad no ha habido fallas durante un tiempo τ. Por la fórmula de la probabilidad condicional

$$h(\tau)\Delta\tau = \frac{w(\lambda, \tau)\Delta\tau}{1 - F(\lambda, \tau)}, \tag{5-122}$$

de donde hallamos que

$$h(\tau) = \frac{w(\lambda, \tau)}{1 - F(\lambda, \tau)}. \tag{5-123}$$

Para la distribución exponencial tenemos:

$$h(\tau) = \frac{\lambda e^{-\lambda\tau}}{e^{-\lambda\tau}} = \lambda, \tag{5-124}$$

es decir, la función del peligro es constante e igual a la intensidad de fallas.

La experiencia de explotación de equipos radioelectrónicos demuestra que la gráfica de la función $h(\tau)$ tiene el aspecto de la

curva representada en la figura 5-8. La sección inicial de esta curva se caracteriza por una elevada intensidad de fallas que se explica por la existencia en algunos elementos de los defectos ocultos que se ponen de manifiesto durante el período inicial de funcionamiento. La intensidad de fallas también crece con grandes τ, lo que es debido al "envejecimiento" de los elementos que se manifiesta en que empeoran las propiedades de éstos después de una prolongada explotación. La intensidad constante de las fallas y, vale decir, la distribución exponencial sólo existe en la sección media de la gráfica $h(\tau)$.

No obstante, es posible librarse de la elevada intensidad de las fallas en el período inicial, sometiendo previamente el equipo a "entrenamiento" antes de comenzar su explotación. Por otro lado, el crecimiento ininterrumpido de los requerimientos de calidad del equipo, resultado del progreso técnico, conduce a que con el tiempo, la calidad de éste deja de satisfacer las nuevas exigencias más

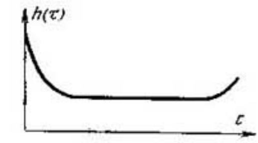


Fig. 5-8 Aspecto típico de la distribución del tiempo hasta la falla

elevadas. Ocurre el gasto moral del equipo que generalmente comienza antes que se inicie un envejecimiento manifiesto de los diferentes elementos. Esto sirve de fundamento en muchos casos para considerar que la dependencia $h(\tau)$ es constante e igual a λ durante todo el período de explotación del equipo, o sea, para emplear la ley exponencial de distribución al evaluar la fiabilidad.

e) Cadenas de Márkov

Sea $Z = \{z_1, \ldots, z_L\}$ el espacio de resultados del experimento o el espacio de estados de cierto sistema, igual para cada paso del proceso estocástico. En tales casos es cómodo designar los estados del sistema simplemente por los subíndices de z, entendiendo por j el estado $z_j \in Z$. Con esto, el espacio de estados va a representarse mediante el conjunto de subíndices $J = \{1, \ldots, L\}$.

Designemos por $\pi_n = (\pi_n^{(1)}, \ldots, \pi_n^{(L)})$ la distribución de probabilidades en el conjunto J para el n-ésimo paso. Entonces, $\pi_n^{(j)}$ determina la probabilidad de que en el n-ésimo paso el sistema se halle en el estado j. Como ya se ha indicado, el proceso estocástico constituye una cadena homogénea de Márkov, si las probabilidades de transición del sistema p_{ij} del estado i al estado j sólo dependen del estado i en el paso anterior y son iguales para cualquier paso.

Las probabilidades de transición pueden representarse por dos métodos diferentes. El primero consiste en anotar las probabilidades de transición en forma de la tabla denominada matriz de transición y designada por **P**. Para $L = 3$ dicha matriz tiene el aspecto:

$$\mathbf{P} \begin{Vmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{Vmatrix}. \qquad (5\text{-}125)$$

La condición (5-108) a que están obligadas a satisfacer las probabilidades de transiciones significa que la suma de los elementos de cualquier línea en la matriz de transiciones deberá ser sin falta igual a la unidad.

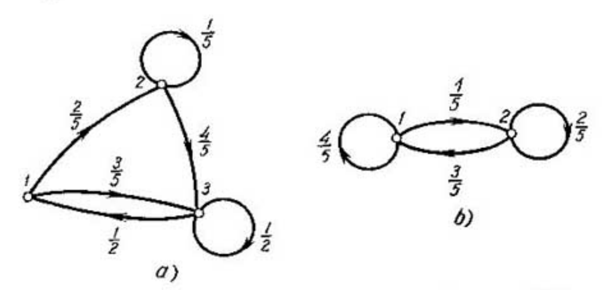


Fig. 5-9. Diagramas de transiciones para la cadena de Márkov

El segundo método para representar las probabilidades de transición consiste en construir el diagrama de transiciones cuyo ejemplo para un sistema con tres estados se muestra en la figura 5-9, *a*. El diagrama de transiciones constituye un grafo cuyos vértices corresponden a los estados del sistema y los arcos direccionales indican las transiciones posibles de un estado a otro. Las probabilidades de transiciones se marcan con los números atribuidos a cada arco. De acuerdo con la condición (5-108) la suma de las probabilidades para los arcos que parten de cualquier vértice del grafo debe ser igual a la unidad.

En distintas ramas de la ciencia y la técnica muchos procesos pueden ser reducidos a las cadenas de Márkov. En sociología las cadenas de Márkov ayudan a estudiar los problemas del cambio de la estructura social o profesional de la población, de la migración de la población, etc. En biología con ayuda de las cadenas de Márkov se estudia la naturaleza del desarrollo de algunas especies de animales y plantas. En física se emplean las cadenas de Márkov para estudiar la difusión de los gases. En técnica con el concurso de las cadenas de Márkov se describen ciertos procesos de transmi-

sión de mensajes, diversos procesos tecnológicos, procesos de control de la capacidad de funcionamiento y de la búsqueda de fallas en sistemas técnicos complejos, etc.

Ejemplo 5-9. Una fábrica produce televisores de una marca determinada. Dependiendo de si el tipo dado de televisor goza o no de demanda entre el público, la fábrica puede encontrarse al final de cada año en uno de dos estados: 1, hay demanda; 2, no hay demanda. Al pasar el tiempo la demanda cambia de tal modo que existe la probabilidad $^4/_5$ de que a finales del año la fábrica se quede en el estado 1. Por otra parte, si resulta que la fábrica se encuentra en el estado 2 se tomarán medidas para modificar y perfeccionar el modelo en producción, así que con la probabilidad $^3/_5$ la fábrica pasará a finales del año siguiente al estado 1.

En el ejemplo dado el desarrollo de la producción se representa por medio de la cadena de Márkov con matriz de transiciones

$$\mathbf{P} \begin{Vmatrix} \frac{4}{5} & \frac{1}{5} \\ \frac{3}{5} & \frac{2}{5} \end{Vmatrix}.$$

El diagrama de transiciones se muestra en la figura 5-9, *b*.

Al estudiar las cadenas de Márkov es necesario aclarar ante todo cómo va a variar la distribución de probabilidades de los estados π_n cuando el sistema adelanta un paso. Designemos por T el suceso que consiste en que en el paso $(n+1)$ el sistema se halla en el estado j.

En correspondencia con las designaciones convenidas la probabilidad de este suceso será igual a $P(T) = \pi_{n+1}^{(j)}$. Designemos por S_i el suceso consistente en que en el n-ésimo paso el sistema se halla en el estado i, de suerte que $P(S_i) = \pi_n^{(i)}$. Con esto, p_{ij} va a representar la probabilidad de que el sistema en el paso $(n+1)$ se encuentre en el estado j si el mismo se hallaba en el estado i en el paso n, es decir, $p_{ij} = P(T|S_i)$. Por la fórmula de la probabilidad completa (5-42) hallamos:

$$\pi_{n+1}^{(j)} = \sum_{i=1}^{L} \pi_n^{(i)} p_{ij}, \qquad j = 1, \ldots, L. \tag{5-126}$$

La fórmula (5-126) permite determinar sucesivamente, paso a paso, el cambio de la distribución de probabilidades de los estados del sistema si se conoce la distribución inicial de probabilidades.

Ejemplo 5-10. Supongamos que en el momento inicial la fábrica del ejemplo 5-9 se halla en el estado 1, o sea, la distribución inicial de probabilidades tiene el aspecto $\pi_0 = (1, 0)$. Utilizando la fórmula (5-126) y la matriz de transiciones, hallamos para el primer paso:

$$\pi_1^{(1)} = \pi_0^{(1)} p_{11} + \pi_0^{(2)} p_{21} = 1 \cdot \frac{4}{5} + 0 \cdot \frac{3}{5} = 0{,}8;$$

$$\pi_1^{(2)} = \pi_0^{(1)} p_{12} + \pi_0^{(2)} p_{22} = 1 \cdot \frac{1}{5} + 0 \cdot \frac{2}{5} = 0{,}2.$$

Para el segundo paso

$$\pi_2^{(1)} = 0{,}76, \qquad \pi_2^{(2)} = 0{,}24.$$

Continuando estos cálculos obtenemos el cambio consecutivo de distribución de las probabilidades dado en la tabla 5-1 con el estado inicial (1,0). En esta tabla se ve que al crecer n la magnitud $\pi_n^{(1)}$ se aproxima a 0,75 y la $\pi_n^{(2)}$, a 0,25

Si el estado inicial del sistema es el estado 2, o sea, $\pi_0 = (0,1)$, la distribución de probabilidades en los pasos consecutivos tendrá el aspecto dado en la tabla 5-1. En este caso al crecer n las magnitudes $\pi_n^{(1)}$ y $\pi_n^{(2)}$ se aproximan a aquellos mismos valores 0,75 y 0,25 que se tenían con la distribución inicial $\pi_0 = (1,0)$. Como vemos, en la cadena de Márkov dada después de un gran número de pasos las probabilidades de transiciones se vuelven independientes del estado inicial del sistema.

Si en una cadena de Márkov existe la distribución límite de las probabilidades correspondiente a $n \to \infty$ e independiente del estado inicial del sistema, entonces esta distribución de probabilidades determina el régimen límite o *estacionario* del sistema. En este caso el sistema se denomina *estáticamente estable* (isostático) y el proceso de Márkov en dicho sistema, *ergódico*.

Designemos por $\pi = (\pi^{(1)}, \ldots, \pi^{L})$ la distribución estacionaria de las probabilidades de una cadena ergódica de Márkov. Las componentes $\pi^{(j)}$ de esta distribución pueden hallarse por las ecuaciones (5-126) que con $n \to \infty$ adquieren la forma:

$$\pi^{(j)} = \sum_{i=1}^{L} \pi^{(i)} p_{ij}, \quad j = 1, \ldots, L. \quad (5\text{-}127)$$

Sin embargo, no todas las L ecuaciónes (5-127) son linealmente independientes, ya que las probabilidades $\pi^{(j)}$ están asociadas entre sí por medio de la correlación

$$\sum_{j=1}^{L} \pi^{(j)} = 1. \qquad (5\text{-}128)$$

Tabla 5-1

Cambio de distribución de probabilidades de los estados del proceso de Márkov

n	Estado inicial $\pi_0=(1,0)$							Estado inicial $\pi_0=(0,1)$						
	0	1	2	3	4	5	...	0	1	2	3	4	5	...
$\pi_n^{(1)}$	1	0,8	0,76	0,752	0,7504	0,75008	...	0	0,6	0,72	0,744	0,7488	0,74976	...
$\pi_n^{(2)}$	1	0,2	0,24	0,248	0,2498	0,24992	...	1	0,4	0,28	0,256	0,2512	0,25024	...

Por eso, para determinar las L componentes desconocidas $\pi^{(j)}$ de la distribución de probabilidades π es suficiente tomar $L-1$ ecuaciones cualesquiera (5-127) y resolverlas junto con la ecuación (5-128).

Ejemplo 5-11. Para la matriz de transiciones en el ejemplo 5-9, la primera de las ecuaciones (5-127) tiene la forma:

$$\pi^{(1)} = \pi^{(1)}p_{11} + \pi^{(2)}p_{21} = \frac{4}{5}\pi^{(1)} + \frac{3}{5}\pi^{(2)}.$$

La ecuación (5-128) da:

$$\pi^{(1)} + \pi^{(2)} = 1.$$

Resolviendo conjuntamente estas ecuaciones, hallamos:

$$\pi^{(1)} = 0{,}75; \qquad \pi^{(2)} = 0{,}25.$$

Se denomina *proceso transitorio* el proceso de transición de la cadena ergódica de Márkov desde su estado inicial hasta el régimen estacionario. El proceso transitorio se describe por la secuencia de distribuciones de las probabilidades π_n para $n = 1, 2 \ldots$ y, con la distribución inicial de probabilidades dada, π_0 puede obtenerse mediante el empleo sucesivo de la fórmula (5-126) como se hizo en el ejemplo 5-8. No obstante, es posible proceder de otra manera.

Introduzcamos en el análisis las magnitudes $p_{ij}^{(l)}$ que determinan la probabilidad de transición del sistema en l pasos del estado i al estado j:

$$p_{ij}^{(l)} = p\left(x_{n+l} = j \mid x_n = i\right), \qquad n = 0,\ 1,\ \ldots \tag{5-129}$$

Las probabilidades $p_{ij}^{(l)}$ hacen posible determinar la distribución de probabilidades π_l por la distribución π_0 según la fórmula análoga a (5-126):

$$\pi_l^{(j)} = \sum_{i=1}^{L} \pi_0^{(i)} p_{ij}^{(l)}, \qquad j = 1,\ \ldots,\ L. \tag{5-130}$$

Para establecer el vínculo de las probabilidades de transición $p_{ij}^{(l)}$ con las probabilidades p_{ij} hagamos $l = l_1 + l_2$, considerando la transición del sistema del estado π_0 al estado π_l en dos etapas: transición del estado π_0 al estado π_{l_1} y después, la transición del estado π_{l_1} al estado $\pi_l = \pi_{l_1+l_2}$. Entonces,

$$\pi_l^{(j)} = \sum_{k=1}^{L} \pi_{l_1}^{(k)} p_{kj}^{(l_2)}, \tag{5-131}$$

donde

$$\pi_{l_1}^{(k)} = \sum_{i=1}^{L} \pi_0^{(i)} p_{ik}^{(l_1)}. \tag{5-132}$$

Sustituyendo el valor de $\pi_{l_1}^{(k)}$ en (5—131), obtenemos:

$$\pi_l^{(l)} = \sum_{k=1}^{L} p_{kl}^{(l_2)} \sum_{i=1}^{L} \pi_0^{(i)} p_{ik}^{(l_1)} = \sum_{i=1}^{L} \pi_0^{(i)} \sum_{k=1}^{L} p_{ik}^{(l_1)} p_{kl}^{(l_2)}. \qquad (5\text{-}133)$$

Comparando (5-133) con (5-130), hallamos que

$$p_{ij}^{(l_1+l_2)} = \sum_{k=1}^{L} p_{ik}^{(l_1)} p_{kj}^{(l_2)}. \qquad (5\text{-}134)$$

Para calcular sucesivamente las probabilidades de transición $p_{ij}^{(1)}$, $p_{ij}^{(2)}$ ... es conveniente hacer $l_1 = l - 1$ y $l_2 = 1$ en (5-134). Con esto obtenemos

$$p_{ij}^{(l)} = \sum_{k=1}^{L} p_{ik}^{(l-1)} p_{kl}. \qquad (5\text{-}135)$$

Suponiendo consecutivamente que $l = 1, 2 \ldots$ y teniendo en cuenta que según la definición de probabilidades de transición p_{ij}

$$p_{ik}^{(0)} = \begin{cases} 1, & i = k; \\ 0, & i \neq k, \end{cases} \qquad (5\text{-}136)$$

obtenemos:

$$p_{ij}^{(1)} = p_{ij}, \quad p_{ij}^{(2)} = \sum_{k=1}^{L} p_{ik} p_{kj}, \ \ldots \qquad (5\text{-}137)$$

La expresión (5-135) puede generalizarse para el caso de una cadena de Márkov no homogénea en la cual las probabilidades de transición dependen del número del paso. En este caso, vamos a entender por probabilidad $p_{ij}(n, n_1)$ para $n < n_1$ la probabilidad de que en el paso n_1 el sistema se encuentre en el estado j si en el paso n éste se encontraba en el estado i:

$$p_{ij}(n, n_1) = p(x_{n_1} = j \mid x_n = i), \quad n < n_1. \qquad (5\text{-}138)$$

Entonces, por analogía con (5-130), tenemos:

$$\pi_{n_1}^{(j)} = \sum_{i=1}^{L} \pi_n^{(i)} p_{ij}(n, n_1), \quad n < n_1. \qquad (5\text{-}139)$$

Por otro lado, considerando cierta n' que satisface la condición $n < n' < n_1$, podemos representar $\pi_{n_1}^{(j)}$ en la forma:

$$\pi_{n_1}^{(j)} = \sum_{k=1}^{L} \pi_{n'}^{(k)} p_{kj}(n', n_1) = \sum_{k=1}^{L} p_{kj}(n', n_1) \sum_{i=1}^{L} \pi_n^{(i)} p_{ik}(n, n') =$$

$$= \sum_{i=1}^{L} \pi_{(n)}^{(i)} \sum_{k=1}^{L} p_{ik}(n, n') p_{kj}(n', n_1). \qquad (5\text{-}140)$$

Comparando (5-139) con (5-140) hallamos:

$$p_{ij}(n, n_1) = \sum_{k=1}^{L} p_{ik}(n, n') p_{kj}(n', n) \qquad (5\text{-}141)$$

para cualquier $n < n' < n_1$. La correlación (5-141) se denomina *ecuación de Chapman — Kolmogórov.*

5-7. ELEMENTOS DE ESTADISTICA MATEMATICA

a) Objeto de la estadística matemática

El desarrollo de la ciencia y la técnica exige ahondar más y más en la esencia de los fenómenos de la naturaleza. Sin embargo, los propios fenómenos de la naturaleza se alzan ante nosotros en forma de una enorme cantidad de hechos y observaciones variados, que a su vez constituyen el resultado de la acción de un conjunto de factores, una parte de los cuales constituye la base del fenómeno considerado, y otros son secundarios, sin importancia y con frecuencia enmascaran sencillamente la esencia del fenómeno. Hacen falta grandes conocimientos y maestría para eliminar toda la información de segundo orden y revelar los datos fundamentales y esenciales contenidos en las observaciones.

Los métodos de estadística matemática brindan la posibilidad de presentar el conjunto de resultados de la observación en una forma compacta y adecuada para su examen. Ellos permiten separar del conjunto de observaciones la información importante presentándola en forma de un pequeño número de índices de resumen. Si resulta que los datos existentes son insuficientes para comprender la esencia del fenómeno y se requiere llevar a cabo un experimento adicional, los métodos de estadística matemática permiten responder a la pregunta de cómo efectuar dicho experimento para simplificar en grado máximo el trabajo del investigador tanto en la realización del experimento como en la elaboración de los datos experimentales.

De lo dicho se infiere que la estadística matemática es la ciencia sobre los métodos para elaborar gran cantidad de datos experimentales con el fin de obtener de ellos deducciones correctas.

En la sección dada no es posible abarcar todos los problemas solucionables por métodos de estadística matemática a los cuales está dedicada toda una serie de libros de texto y monografías. Por eso, sólo nos limitaremos a estudiar los métodos más difundidos para resolver los problemas estadísticos. En el capítulo 9 se ofrecerá el desarrollo de algunos de estos métodos realizado desde el punto de vista de la teoría de toma de decisiones.

b) Concepto de muestreo aleatorio

Sea x una magnitud aleatoria unidimensional con la función de distribución de probabilidades $F(x)$. Examinemos la magnitud aleatoria de n dimensiones

$$(x^{(1)}, \ldots, x^{(n)}), \tag{5-142}$$

cuyas componentes separadas son magnitudes aleatorias independientes con iguales funciones de distribución de probabilidades $F(x)$. La función de distribución de probabilidades de esta magnitud multidimensional se determina como el producto de las funciones de distribución de probabilidades de sus componentes

$$\prod_{i=1}^{n} F(x^{(i)}). \tag{5-143}$$

Es cómodo considerar las magnitudes aleatorias $x^{(i)}$ como resultados de cierto experimento. Entonces la magnitud aleatoria multidimensional (5-142) puede considerarse como el resultado de la realización consecutiva de n experimentos independientes en una misma instalación experimental o como resultado de n experimentos simultáneos en n instalaciones experimentales de un mismo tipo.

Ya que siempre existe la posibilidad siquiera teórica de realizar un número ilimitado de experimentos puede hablarse de una colección infinita de números aleatorios $(x^{(i)})$ con la función de distribución de probabilidades $F(x)$. Esta colección infinita se denomina *conjunto infinito con función de distribución de probabilidades $F(x)$*.

Cada resultado concreto del experimento, y por lo tanto, cada $x^{(i)}$ de la colección finita (5-142), puede considerarse como la selección de uno de los números del conjunto infinito. La colección completa (5-142) representa la secuencia de n de estas muestras y se denomina *muestreo aleatorio*.

c) Teoremas límites de la teoría de las probabilidades

Presenta gran interés en la estadística matemática estudiar el cambio de propiedades del muestreo aleatorio al aumentar su volumen y, en particular, las correlaciones límites que se obtienen con $n \to \infty$. Con esto, juegan un papel principalísimo la desigualdad de Chébishev, la ley de los grandes números que se deriva de ella y el teorema central límite.

Si x es una magnitud aleatoria con esperanza matemática ν y dispersión σ^2, es válida la desigualdad siguiente llamada desigualdad de Chébishev:

$$P(|x-\nu| \geqslant \lambda\sigma) \leqslant \frac{1}{\lambda^2}. \tag{5-144}$$

Para demostrar esta desigualdad fraccionemos el eje real en tres intervalos:

$$I = (-\infty, \nu - \lambda\sigma]; \quad I' = (\nu - \lambda\sigma, \nu + \lambda\sigma); \quad I'' = [\nu + \lambda\sigma, +\infty).$$

La dispersión de la magnitud x puede escribirse en la forma:

$$\sigma^2 = \int_{-\infty}^{+\infty} (x - \nu)^2 w(x) dx = \int_I (x - \nu)^2 w(x) dx +$$

$$+ \int_{I'} (x - \nu)^2 w(x) dx + \int_{I''} (x - \nu)^2 w(x) dx.$$

Omitiendo en las expresiones para σ^2 el sumando correspondiente al intervalo I', obtenemos:

$$\sigma^2 \leqslant \int_I (x - \nu)^2 w(x) dx + \int_{I''} (x - \nu)^2 w(x) dx.$$

Por cuanto en el intervalo I $x - \nu \leqslant -\lambda\sigma$ y en el I'' $x - \nu \geqslant \lambda\sigma$ sólo se incrementa la desigualdad si en ambos integrales sustituimos $(x - \nu)^2$ por $\lambda^2\sigma^2$. Así pues,

$$\sigma^2 \leqslant \lambda^2\sigma^2 [P(x \leqslant \nu - \lambda\sigma) + P(x \geqslant \nu + \lambda\sigma)] = \lambda^2\sigma^2 P(|x - \nu| \geqslant \lambda\sigma)$$

lo que equivale a la desigualdad (5-144).

La desigualdad (5-144) determina la probabilidad de que la magnitud $x - \nu$ salga de los límites del intervalo $(-\lambda\sigma, +\lambda\sigma)$. No obstante, puede determinarse la probabilidad de que la magnitud $x - \nu$ no salga de los límites del intervalo indicado. Esta probabilidad es igual a:

$$P(|x - \nu| < \lambda\sigma) = 1 - P(|x - \nu| \geqslant \lambda\sigma),$$

o tomando en cuenta la desigualdad (5-144),

$$P(|x - \nu| < \lambda\sigma) \geqslant 1 - \frac{1}{\lambda^2}. \tag{5-145}$$

Si designemos $\lambda\sigma$ por ε, esta desigualdad toma la forma:

$$P(|x - \nu| < \varepsilon) \geqslant 1 - \frac{\sigma^2}{\varepsilon^2}. \tag{5-146}$$

Vamos a referirnos al análisis del muestreo aleatorio $(x_1, \ldots, x_n)$ de volumen n. Designemos por $\bar{x}$ la media aritmética de este muestreo, que es igual a:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i. \tag{5-147}$$

Con el objeto de simplificar las notaciones de semejante tipo en lo sucesivo vamos a utilizar el símbolo S o S_n para designar la

sumación de todos los datos de muestreo. En tal caso la correlación (5-147) se escribirá en la forma:

$$\bar{x} = \frac{1}{n} S(x) = \frac{S_n}{n}. \quad (5\text{-}148)$$

La dispersión de la magnitud $\bar{x}$ según (5-85) y (5-100) será igual a:

$$D(\bar{x}) = D\left[\frac{S(x)}{n}\right] = \frac{1}{n^2} S[D(x)] = \frac{\sigma^2}{n}. \quad (5\text{-}149)$$

Sustituyendo en (5-146) x por $\bar{x}$ y respectivamente σ^2 por σ^2/n, obtendremos:

$$P\left\{\left|\frac{S_n}{n} - \nu\right| < \varepsilon\right\} \geqslant 1 - \frac{\sigma^2}{n\varepsilon^2}, \quad (5\text{-}150)$$

donde ν es la esperanza matemática $\bar{x}$ que constituye el valor medio para el conjunto infinito del que se ha tomado la muestra.

En la expresión (5-150) se ve que con $n \to \infty$ y cualquier $\varepsilon > 0$

$$P\left\{\left|\frac{S_n}{n} - \nu\right| < \varepsilon\right\} \to 1, \quad (5\text{-}151)$$

lo que expresa la *ley de grandes números*. De ella se deduce que con grandes n y una probabilidad tan cercana a cero como sea necesaria, la media aritmética coincide con la esperanza matemática $\nu = M(x)$ de la magnitud x.

Expongamos sin demostración una de las correlaciones límites más importantes de la teoría de las probabilidades que ha recibido el nombre de *teorema central límite*. Sea $(x_1, \ldots, x_n)$ una secuencia de magnitudes aleatorias que, en el caso general, tienen diferentes funciones de distribución de probabilidades $F_1(x), \ldots, F_n(x)$, los valores medios $\nu_1, \ldots, \nu_n$ y las dispersiones $\sigma_1^2, \ldots, \sigma_n^2$. Designemos

$$\bar{x} = \frac{1}{n} S(x); \quad \nu = \frac{1}{n} S(\nu); \quad s^2 = \frac{1}{n} S(\sigma^2). \quad (5\text{-}152)$$

El teorema central límite afirma que si las magnitudes aleatorias $x_1, \ldots, x_n$ son independientes y tienen dispersiones finitas de un mismo orden, entonces, habiendo un gran número de sumandos, la magnitud de la media aritmética x se aproxima a la distribución normal con el valor medio ν y la dispersión s^2. Observando que en este caso la magnitud

$$t = \frac{\bar{x} - \nu}{s/\sqrt{n}} \quad (5\text{-}153)$$

tendrá valor medio nulo y dispersión unitaria, basándonos en el teorema central límite podemos afirmar que con $n \to \infty$ se cumplirá la correlación

$$P\left(\frac{\bar{x} - \nu}{s/\sqrt{n}} < y\right) \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-\frac{1}{2}u^2} du. \quad (5\text{-}154)$$

Con mucha frecuencia nos tropezamos con problemas en los cuales la magnitud aleatoria a analizar es la suma de un gran número de magnitudes aleatorias independientes. Por ejemplo, el error de un aparato complicado es el resultado de la acción combinada de variadas condiciones exteriores y errores en algunos elementos de este aparato. En tal caso los errores de gran influencia se eliminan con facilidad.

Así pues, si ocurrió una ruptura en un circuito eléctrico, es fácil descubrirla y eliminar. Sin embargo, es difícil encontrar las causas de los errores de pequeña magnitud. Luego, puede considerarse que son uniformemente pequeños los sumandos separados del error suma y afirmarse fundándose en el teorema central límite, que el error suma va a tener una distribución próxima a la normal.

d) Problemas de la estadística matemática

El problema fundamental de estadística matemática es la determinación de los parámetros desconocidos de distribución del conjunto infinito por una muestra finita conocida. Con eso podemos referirnos a determinar la propia función de distribución $F(x)$ o los parámetros separados de la función de distribución tales como el valor medio, la dispersión, el valor mínimo o máximo de una magnitud aleatoria, etc.

Por cuanto los elementos de una muestra finita son magnitudes aleatorias, será también aleatorio el valor del parámetro determinado con ayuda de esta muestra. En particular, si tenemos varias muestras de igual extensión de un mismo conjunto infinito, cada una de ellas dará su valor del parámetro que nos interesa. Es por eso que mediante una muestra finita no podemos juzgar con precisión sobre el valor del parámetro y únicamente podemos estimar con mayor o menor exactitud este parámetro. Los valores numéricos de algunos parámetros determinados de la muestra finita se denominan *estimaciones* de los parámetros del conjunto infinito.

En el caso general, de una misma muestra pueden obtenerse diferentes estimaciones. Supongamos, por ejemplo, que existe la muestra $\{x_1, x_2, \ldots, x_n\}$ siendo $x_1 \leqslant x_2 \leqslant \ldots \leqslant x_n$ y queremos hallar la estimación del valor medio. En concordancia con las fórmulas (3-24), (3-25) y (3-27) podemos examinar varias estimaciones de la media: la media de todas las observaciones denominada *mediana*:

$$\hat{\bar{x}} = \begin{cases} x_{\frac{n+1}{2}} & \text{para } n \text{ impar;} \\ \frac{1}{2}\left(x_{\frac{n}{2}-1} + x_{\frac{n}{2}+1}\right) & \text{para } n \text{ par;} \end{cases} \tag{5-155}$$

la semisuma de los valores mínimo y máximo:

$$\hat{\bar{x}} = \frac{1}{2}(x_1 + x_n); \tag{5-156}$$

la media aritmética de todas las observaciones:

$$\hat{\bar{x}} = \frac{1}{n} \sum_{i=1}^{n} x_i. \qquad (5\text{-}157)$$

¿Cuál de estas estimaciones es preferible? Puede darse respuesta a esta pregunta si se introducen ciertos criterios a los que debe satisfacer la estimación. Examinemos los más importantes de ellos.

1. *No desplazamiento.* La estimación no debe contener error sistemático que aumente o disminuya el valor del parámetro para todas las muestras. Esto significa que la esperanza matemática de la estimación debe coincidir con el valor real del parámetro. Si designamos por α el valor real del parámetro y por $\hat{\alpha}$, su estimación, el requisito de no desplazamiento se escribe en la forma

$$M(\hat{\alpha}) = \alpha. \qquad (5\text{-}158)$$

2. *Validez.* La estimación $\hat{\alpha}$ debe aproximarse a la magnitud del parámetro α a medida que aumente la extensión de la muestra. Como que la estimación $\hat{\alpha}$ es una magnitud aleatoria, sólo se puede hablar de esta aproximación en sentido probabilístico. Así pues, si designamos por $\hat{\alpha}_n$ la estimación α obtenida por la muestra del volumen n, para una estimación válida debe cumplirse la correlación

$$P[|\hat{\alpha}_n - \alpha| < \varepsilon] \to 1 \qquad (5\text{-}159)$$

para $n \to \infty$ y cualquier $\varepsilon > 0$.

3. *Eficacia.* De todas las estimaciones no desplazadas y válidas hay que preferir aquella que resulta más próxima al parámetro estimado, es decir, con la que, al utilizar las diferentes muestras, las desviaciones grandes se encuentren lo más raramente posible. Las estimaciones que satisfacen este requisito se denominan *eficaces*. El requisito matemático de eficacia significa que se requiere la dispersión mínima de la estimación

$$D(\hat{\alpha} - \alpha) = \min. \qquad (5\text{-}160)$$

e) Estimaciones no desplazadas de la media y la dispersión

Sea $(x_1, \ldots, x_n)$ la muestra de un conjunto con el valor medio ν y la dispersión σ^2. Designemos por $\bar{x}$ la estimación para la media y por s^2, la estimación para la dispersión.

Habitualmente, se toma como estimación de la magnitud ν el valor medio aritmético de la selección

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i = \frac{1}{n} S(x), \qquad (5\text{-}161)$$

Valores de límite del 95% de χ_q^2 para la distribución χ^2

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14
χ_q^2	0,004	0,103	0,352	0,71	1,14	1,63	2,17	2,73	3,32	3,94	4.6	5,2	5.9	6,6

denominado *media muestral*. Como se desprende de la ley de los grandes números, tal estimación es válida, o sea, tiende a ν cuando $n \to \infty$.

Por cuanto la media muestral $\bar{x}$ representa una suma de magnitudes aleatorias, dicha media será una magnitud aleatoria. Por eso puede hablarse de la ley de distribución de la media muestral, de su valor medio y la dispersión.

El valor medio o esperanza matemática de la media muestral es igual a:

$$M(\bar{x}) = \frac{1}{n} M[S(x)] = \frac{1}{n} S[M(x)] = \nu. \tag{5-162}$$

Como vemos, la esperanza matemática de la media muestral coincide con el valor ν lo que es indicio del no desplazamiento de la estimación (5-161). La dispersión de la media muestral según (5-149) es igual a σ^2/n.

Si la muestra está tomada del conjunto con distribución normal, entonces $\bar{x}$, como suma de las magnitudes aleatorias normalmente distribuidas, va a tener una distribución normal que, tomando en consideración (5-162) y (5-149), tendrá la forma:

$$w(\bar{x}) = N\left(\nu, \frac{\sigma^2}{n}\right). \tag{5-163}$$

En muchos casos en lugar de $\bar{x}$ es cómodo examinar la magnitud

$$t = \frac{\bar{x} - \nu}{\sigma/\sqrt{n}} = \frac{\sqrt{n}\,(\bar{x} - \nu)}{\sigma}, \tag{5-164}$$

que tiene media nula y la dispersión unitaria, cuya densidad de distribución de probabilidades será igual a $N(0, 1)$.

Si la muestra fue tomada de un conjunto que tiene una distribución distinta de la normal, la ley de distribución de la media muestral también va a diferir de la normal, no obstante, basándose en el teorema central límite, puede considerarse que con grandes volúmenes de muestras esta ley será próxima a la ley $N(\nu, \sigma^2/n)$.

Para encontrar la estimación no desplazada de la dispersión s^2 hallaremos la esperanza matemática de la magnitud $S(x - \bar{x})^2$.

en dependencia del número n de grados de libertad

15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
7,3	8,0	8,7	9,4	10,1	10,9	11,6	12,3	13,1	13,8	14,6	15,4	16,2	16,9	17,7	18,5

Representando esta magnitud en la forma

$$S(x-\bar{x})^2 = S[(x-\nu)-(\bar{x}-\nu)]^2 =$$
$$= S(x-\nu)^2 - 2(\bar{x}-\nu)S(x-\nu) + n(\bar{x}-\nu)^2$$

y tomando la esperanza matemática de cada sumando, obtenemos:

$$M[S(x-\bar{x})^2] = (n-1)\sigma^2, \tag{5-165}$$

o bien

$$M\left[\frac{S(x-\bar{x})^2}{n-1}\right] = \sigma^2. \tag{5-166}$$

De (5-166) se deduce que puede servir de valor no desplazado de la dispersión la estimación de la forma

$$s^2 = \frac{1}{n-1}S(x-\bar{x})^2. \tag{5-167}$$

Para determinar la ley de distribución de la estimación de dispersión examinemos previamente una importante ley de distribución. Si y_i son magnitudes aleatorias independientes que obedecen a la distribución normal $N(0, 1)$, entonces la magnitud aleatoria

$$\chi^2 = \sum_{i=1}^{n} y_i^2 \tag{5-168}$$

obedece a la distribución χ^2 (ji-cuadrado) con n grados de libertad. En este caso el número de grados de libertad se determina por el número de magnitudes aleatorias independientes en la suma (5-168). Para la distribución χ^2 se han compilado tablas detalladas en las cuales se brindan los valores de la probabilidad $p(\chi^2-\chi_q^2)$ para los distintos n, donde χ_q^2 es un número positivo cualquiera.

La tabla 5-2 es un ejemplo de tabla similar en la que se han presentado los valores de χ_q^2 que satisfacen la condición $p(\chi^2-\chi_q^2)=0{,}95$ para n del 1 al 30.

No es difícil ver que si por y_i se entiende la magnitud $(x_i-x)/\sigma$ que satisface las condiciones impuestas a y_i, entonces

la magnitud

$$\chi^2 = \sum_{i=1}^{n} y_i^2 = \frac{(n-1)\,s^2}{\sigma^2} \tag{5-169}$$

va a definir la ley de distribución para s^2.

f) Búsqueda de las estimaciones por el método de máxima verosimilitud

Muchos problemas de estadística matemática se reducen a estimar cierto parámetro α en una distribución cuyo tipo es conocido. Por el principio de máxima verosimilitud, como estimación del parámetro α se toma el valor que parezca el más probable fundándose en los datos experimentales.

Supongamos que tenemos la muestra $(x_1, \ldots, x_n)$ de un conjunto con densidad de distribución de probabilidades $w(x, \alpha)$, que depende del parámetro α. La densidad multidimensional de distribución de probabilidades para esta muestra

$$L = \prod_{i=1}^{n} w(x_i, \alpha) \tag{5-170}$$

se denomina *función de verosimilitud*. Según el principio de máxima verosimilitud, se toma como estimación del parámetro α el valor $\hat{\alpha}$ con el cual la función de verosimilitud L o, lo que es igual, el logaritmo de dicha función $\ln L$ alcanza el máximo. El valor $\hat{\alpha}$ puede hallarse por la ecuación de verosimilitud

$$\frac{\partial \ln L}{\partial \alpha} = \sum_{i=1}^{n} \frac{\partial \ln (x_i, \alpha)}{\partial \alpha} = 0, \tag{5-171}$$

que se reduce fácilmente a la forma

$$\sum_{i=1}^{n} \frac{1}{w(x_i, \alpha)} \frac{\partial w(x_i, \alpha)}{\partial \alpha} = 0. \tag{5-172}$$

Para el caso en que la muestra se ha tomado del conjunto con distribución normal $N(\nu, \sigma^2)$ definida por (5-54), las ecuaciones de verosimilitud para determinar los parámetros ν y σ son de la forma:

$$\left.\begin{aligned} &\sum_{i=1}^{n} (x_i - \nu) = 0; \\ &\sum_{i=1}^{n} \left[1 - \frac{(x_i - \nu)^2}{\sigma^2}\right] = 0. \end{aligned}\right\} \tag{5-173}$$

De la primera ecuación hallamos la estimación para v

$$\hat{v} = \frac{1}{n} \sum_{i=1}^{n} x_i = \bar{x}, \tag{5-174}$$

que coincide con la estimación no desplazada y válida (5-161) obtenida antes. La segunda ecuación de verosimilitud tomando en cuenta (5-174) da para σ^2 la estimación,

$$S^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2, \tag{5-175}$$

que, sin embargo, es desplazada. Por eso, habitualmente en la práctica se emplea la estimación no desplazada (5-167).

Examinemos en el ejemplo siguiente cómo se obtiene la estimación de verosimilitud máxima para la distribución exponencial.

Ejemplo 5-12. Para determinar el plazo medio de duración del equipo producido por una fábrica, se probaron n muestras (modelos) del equipo durante t horas. En el transcurso de la prueba m modelos funcionaron sin fallas durante todo el tiempo t, y $n-m$ modelos fallaron en los momentos $\tau_1, \ldots, \tau_{n-m}$. ¿Cuál es la estimación más verosímil del período medio de funcionamiento sin fallas?

Suponiendo que el tiempo en el cual ocurre la falla del equipo obedece a la ley exponencial (5-120) de tal suerte que la magnitud $w(\lambda, \tau) = \lambda e^{-}$ representa la probabilidad de que el equipo se averie en el momento τ y la magnitud $e^{-\lambda t}$, la probabilidad de que funcione normalmente durante el tiempo t, obtenemos la función de verosimilitud en la forma

$$L = (e^{-\lambda t})^m \prod_{i=1}^{n-m} w(\lambda, \tau_i). \tag{5-176}$$

Con esto la ecuación (5-171) da:

$$-mt + \sum_{i=1}^{n-m} \left(\frac{1}{\lambda} - \tau_i\right) = 0, \tag{5-177}$$

de donde

$$\tau_{\text{med}} = \frac{1}{\lambda} = \frac{mt + \sum_{i=1}^{n-m} \tau_i}{n-m}. \tag{5-178}$$

De manera análoga puede mostrarse que la estimación más verosímil de la probabilidad de un suceso para el proceso de pruebas independientes, en caso de distribución binominal, es igual a:

$$\hat{p} = \frac{x}{n}, \tag{5-179}$$

donde n es el número de pruebas y x, el de éxitos.

Si el proceso aleatorio obedece a la distribución de Poisson y en la muestra de n intervalos están contenidos r sucesos, entonces sirve de estimación más verosímil del número medio de sucesos en un intervalo la magnitud

$$\hat{a} = \frac{r}{n}. \tag{5-180}$$

g) Estimación de los parámetros por el método de intervalos de confianza

En las secciones precedentes se estudiaron los métodos que brindan la estimación del parámetro en forma de cierto número, o sea, de cierto punto del eje real. Tales estimaciones recibieron la denominación de *puntuales*. Por cuanto la muestra es aleatoria, en muchos casos, sobre todo con pequeños volúmenes de muestra, esta estimación puede resultar muy alejada del valor real del parámetro.

Puede proporcionar mucha más información el indicar el intervalo en el cual se halla con probabilidad próxima a la unidad, el valor del parámetro a estimar. De esta manera, si γ es una magnitud próxima a la unidad, el intervalo (α_1, α_2) en el que con la probabilidad γ se encuentra el valor del parámetro desconocido α

$$P(\alpha_1 < \alpha < \alpha_2) = \gamma \tag{5-181}$$

se denomina intervalo de confianza del 100 γ %. En la práctica, a menudo nos limitamos al examen del intervalo de confianza de un 95% suponiendo que $\gamma = 0{,}95$. Si se ha hallado la estimación puntual del parámetro $\hat{\alpha}$ y se conoce la densidad de distribución de probabilidades para esta estimación, los límites del intervalo de confianza se obtienen por la correlación que sigue:

$$\int_{\alpha_1}^{\alpha_2} w(\hat{\alpha})\, d\hat{\alpha} = \gamma. \tag{5-182}$$

En diversos casos en vez de hallar el intervalo bilateral de confianza (α_1, α_2) sólo se requiere hallar el límite inferior α_1' o el superior α_2' para el parámetro considerado α. Los intervalos definidos por las correlaciones

$$P(\alpha > \alpha_1') = \gamma \text{ y } P(\alpha < \alpha_2') = \gamma, \tag{5-183}$$

se denominan *intervalos unilaterales de confianza del* 100 γ *por ciento*. Los límites α_1' y α_2' de estos intervalos se determinan por las expresiones

$$\int_{\alpha_1'}^{+\infty} w(\hat{\alpha})\, d\hat{\alpha} = \gamma \text{ y } \int_{-\infty}^{+\alpha_2'} w(\hat{\alpha})\, d\hat{\alpha} = \gamma. \tag{5-184}$$

Supongamos que la muestra del volumen n está tomada del conjunto con distribución normal $N(\nu, \sigma^2)$ con la dispersión desconocida ν y conocida σ^2 de suerte que $t = \sqrt{n}(\bar{x} - \nu)/\sigma$ va a tener la distribución $N(0, 1)$. Introduzcamos asimismo la designación $\gamma = 1 - q$. En este caso el intervalo de porcentaje de confianza

$100(1-q)$ se determina mediante la correlación

$$P\left(-t_q < \frac{\sqrt{n}\,(\bar{x}-\nu)}{\sigma} < +t_q\right) = 1-q, \qquad (5\text{-}185)$$

o bien

$$P\left(\bar{x} - t_q\frac{\sigma}{\sqrt{n}} < \nu < \bar{x} + t_q\frac{\sigma}{\sqrt{n}}\right) = 1-q. \qquad (5\text{-}186)$$

Teniendo en cuenta (5-58) esta expresión puede escribirse en la forma

$$1-q = \Phi(t_q) - \Phi(-t_q) = 2\Phi(t_q), \qquad (5\text{-}187)$$

lo que conduce a la siguiente correlación para determinar t_q:

$$\Phi(t_q) = \frac{1}{2}(1-q). \qquad (5\text{-}188)$$

Para los intervalos unilaterales de confianza $(t'_q, +\infty)$ y $(-\infty, t'_q)$ la correlación (5-187) se escribe en la forma

$$1-q = \Phi(t'_q) - \Phi(-\infty) = \frac{1}{2} + \Phi(t'_q). \qquad (5\text{-}189)$$

Con esto, para determinar los valores límites de los intervalos unilaterales de confianza se obtiene la correlación

$$\Phi(t'_q) = \frac{1}{2}(1-2q). \qquad (5\text{-}190)$$

Como vemos, los valores límites del $100(1-q)$ por ciento del intervalo unilateral de confianza son iguales a los valores límites del intervalo bilateral de confianza del $100(1-2q)$ por ciento.

Haciendo $q = 0{,}05$ y utilizando las tablas de la integral de probabilidades de las correlaciones (5-188) y (5-190) encontramos los siguientes límites de los intervalos bilaterales y unilaterales de confianza:

$$t_q = 1{,}96; \quad t'_q = 1{,}64. \qquad (5\text{-}191)$$

De este modo, la media muestral $\bar{x}$ calculada por la muestra del volumen n del conjunto con distribución normal y dispersión conocida σ^2 hace posible afirmar con una probabilidad de 0,95 que el valor medio ν del conjunto se determinará en el intervalo

$$\bar{x} - 1{,}96\frac{\sigma}{\sqrt{n}} < \nu < \bar{x} + 1{,}96\frac{\sigma}{\sqrt{n}}, \qquad (5\text{-}192)$$

o que

$$\nu > \bar{x} - 1{,}64\frac{\sigma}{\sqrt{n}}, \qquad (5\text{-}193)$$

o

$$\nu < \bar{x} + 1{,}64\frac{\sigma}{\sqrt{n}}. \qquad (5\text{-}194)$$

Los valores obtenidos t_q pueden utilizarse asimismo para buscar los intervalos de confianza con la distribución binomial

$w(x, n, p)$ por cuanto dicha distribución con grandes n se aproxima a la normal con el valor medio $\nu = np$ y la dispersión $\sigma^2 = npq$. Haciendo $t = (x - np)/\sqrt{npq}$, llegamos a la conclusión de que si durante n pruebas independientes el suceso S ocurrió x veces, puede afirmarse con una probabilidad de 0,95 que el valor medio np se hallará entre los límites siguientes:

$$x - 1{,}96\sqrt{npq} < np < x + 1{,}96\sqrt{npq}, \qquad (5\text{-}195)$$

o que

$$np > x - 1{,}64\sqrt{npq}, \qquad (5\text{-}196)$$

o bien

$$np < x + 1{,}64\sqrt{npq}, \qquad (5\text{-}197)$$

con esto, la magnitud $\sqrt{npq}$ se calcula por el valor más verosímil $p = \hat{p} = x/n$ como

$$\sqrt{npq} = \sqrt{\frac{x(n-x)}{n}}. \qquad (5\text{-}198)$$

Ejemplo 5-13. De la producción elaborada por una fábrica se eligieron y probaron aleatoriamente 100 artículos entre los cuales 37 resultaron de segunda clase. ¿Qué puede decirse del porcentaje de artículos de segunda clase producidos?

En el caso dado $n = 100$, $x = 37$, $\sqrt{npq} = 4{,}84$. Con esto $1{,}93\sqrt{npq} = 9{,}5$ de manera que $x - 1{,}96\sqrt{npq} = 27{,}5$ y $x + 1{,}96\sqrt{npq} = 46{,}5$.

De este modo hay 5 posibilidades de 100 de que nos equivoquemos haciendo una de las afirmaciones siguientes respecto al número de artículos de segunda clase elaborados:

1) el mismo está comprendido entre el 27,5 y el 46,5% del número total de artículos;

2) no es menor del 29% del número total de artículos;

3) no es mayor del 45% del número total de artículos.

Con gran frecuencia es necesario utilizar los intervalos de confianza para determinar la estimación ν con la distribución normal cuando se desconoce la dispersión σ^2. En este caso en lugar de σ^2 resulta necesario utilizar la estimación de la dispersión s^2. Pero entonces la magnitud

$$t = \frac{\sqrt{n}\,(\bar{x} - \nu)}{s}, \qquad (5\text{-}199)$$

a diferencia de (5-164), ya no estará distribuida según la ley normal. En realidad, la magnitud t definida por (5-199) va a obedecer a la ley de distribución de Student, a la que puede darse la definición siguiente.

Sea u la magnitud aleatoria que tiene la distribución normal $N(0, 1)$ y v, la que posee la distribución χ^2. Si u y v son independientes, la magnitud aleatoria

$$t = \frac{u}{\sqrt{v/k}} = \frac{u\sqrt{k}}{\sqrt{v}} \qquad (5\text{-}200)$$

define la *distribución de Student* en la cual la magnitud k se denomina número de grados de libertad.

Si determinamos ahora las magnitudes

$$u = \frac{\sqrt{n}\,(\bar{x} - \nu)}{\sigma}; \quad v = \frac{(n-1)\,s^2}{\sigma^2}, \tag{5-201}$$

según (5-164) y (5-169) éstas van a satisfacer las condiciones que definen la distribución de Student, y la magnitud t, determinada según (5-200) coincidirá con el valor estipulado por (5-199) para $k = n - 1$. De este modo, la magnitud aleatoria t definida por la correlación (5-199) estará distribuida según la ley de Student con $k = n - 1$ grados de libertad.

Designemos por t_{qk} los límites del intervalo bilateral de confianza del $100(1-q)$ por ciento para la distribución de Student con k grados de libertad. Existen tablas detalladas para determinar t_{qk} con distintas q para k de 1 a 30. Con $k > 30$ los valores de t_{qk} coinciden prácticamente con los valores de t_q obtenidos para la distribución normal. El ejemplo de dicha tabla es la tabla 5-3 que proporciona los valores de t_{qk} para $q = 0{,}05$ y $0{,}1$.

Los límites de los intervalos unilaterales de confianza t'_{qk} pueden hallarse por los valores t_{qk} mediante la correlación

$$t'_{qk} = t_{2qk}. \tag{5-202}$$

Ejemplo 5-14. Para disminuir la influencia de las interferencias en el canal de comunicación cada resultado de la medición de cierto parámetro ν se transmite tres veces desde a bordo de un cohete a la estación de medida en tierra. Los resultados de las medidas son iguales a $x_1 = 3{,}2$; $x_2 = 2{,}9$ y $x_3 = 3{,}1$. Determinar el intervalo bilateral de confianza del 95% para el parámetro ν considerando las distorsiones en cada valor transmitido independientes unas de otras y destribuidas según la ley normal.

Tabla 5-3

Valores de los límites del intervalo de confianza del 100 $(1-q)$ por ciento en dependencia del número de grados de libertad k para la distribución de Student

q	k																				
	1	2	3	4	5	6	7	8	9	10	12	14	16	18	20	22	24	26	28	30	∞
0,05	12,71	4,30	3,18	2,78	2,57	2,45	2,36	2,31	2,62	2,23	2,18	2,15	2,12	2,10	2,09	2,07	2,06	2,06	2,05	2,04	1,96
0,10	6,31	2,92	2,35	2,13	2,01	1,94	1,89	1,86	1,83	1,81	1,78	1,76	1,75	1,73	1,72	1,72	1,71	1,71	1,70	1,70	1,64

Por los datos medidos hallamos que

$$n = 3, \quad \bar{x} = \frac{1}{n} S(x) = 3{,}067;$$

$$s^2 = \frac{1}{n-1} S(x - \bar{x})^2 = 0{,}0234.$$

Por la tabla 5-3 para $q = 0{,}05$ y $k = n - 1 = 2$, obtenemos:

$$t_{qk} = 4{,}30, \text{ así que } t_{qk} = s/\sqrt{n} = 0{,}377.$$

Luego, con la probabilidad de 0,95 tenemos que $2{,}69 < v < 3{,}44$.

h) Verificación de las hipótesis estadísticas. Concepto de criterio de acuerdo

Sea Z el espacio de resultados del experimento con los elementos $z \in Z$. Efectuemos n experimentos para determinar la probabilidad $p(z)$ de cierto resultado z. Si con esto el resultado z ocurrió n_z veces, entonces estimamos la probabilidad de dicho resultado por la relación n_z/n. No obstante, en realidad la relación n_z/n no nos da la probabilidad $p(z)$ del resultado z, sino su frecuencia $q(z)$ que puede diferir considerablemente de la probabilidad, sobre todo con un número no muy grande de experimentos. Surge la pregunta, en qué medida podemos juzgar sobre la distribución de probabilidades de los resultados $p(z)$ si conocemos las frecuencias de estos resultados $q(z)$. Sin embargo, este planteo del problema no es del todo correcto.

Efectivamente, si a consecuencia del experimento se ha obtenido una serie de números que expresan las frecuencias de los resultados aislados del experimento $q(z)$, necesitamos determinadas razones para cambiarlos por otros números [probabilidades $p(z)$]. ¿Qué puede servir de fundamento para sustituir unos números por otros?

Supongamos que al lanzar 10 veces una moneda ésta cayó de escudo 7 veces. La frecuencia de caída de escudo en el experimento dado es de 0,7. ¿Acaso tenemos fundamento para cambiar esta magnitud y tomar como probabilidad de caída de escudo algún otro número?

Si no sabemos nada de la moneda, en particular, si es simétrica o no, tampoco tenemos razón para cambiar el valor de la frecuencia obtenido del experimento. No obstante, el problema puede plantearse de otra manera. Sabemos que si la moneda es simétrica, la probabilidad de que caiga de escudo es igual a 0,5. Queremos comprobar si es simétrica la moneda y con esta finalidad la lanzamos 10 veces. Al hacerlo, ella cae de escudo 7 veces. Nos preguntamos: ¿tenemos o no fundamento para afirmar que la moneda es simétrica?

El problema así formulada se denomina *problema de verificación de la hipótesis*. Problemas similares surgen cuando existe fundamento para construir cierta hipótesis (por ejemplo, la moneda es

simétrica) sobre el carácter de la distribución o del valor de los parámetros de distribución de la magnitud aleatoria. El objetivo del experimento es confirmar o refutar la hipótesis propuesta. Los métodos de verificación estadística de las hipótesis permiten responder a las preguntas:

1) ¿tiene o no el nuevo modelo de aparato calidad superior en comparación con los existentes?

2) ¿va a poseer o no el sistema una probabilidad no inferior a la deseada?

3) ¿va a ser o no el nuevo método de curación más eficaz que los existentes? etc.

Para tener derecho a aceptar o rechazar la hipótesis considerada es necesario elaborar cierto criterio que se denomina *criterio de acuerdo* de la hipótesis verificada con los resultados del experimento.

Supongamos que basándonos en razones cualesquiera consideramos que el valor de cierto parámetro α es igual a α_0 y queremos comprobarlo fundándonos en un experimento. Como resultado del experimento hallamos la estimación de este parámetro $\hat{\alpha}$ que, siendo una magnitud aleatoria, no coincide por lo general con el valor α_0. Pero, la diferencia $\Delta\alpha = \hat{\alpha} - \alpha_0$ de la estimación $\hat{\alpha}$ respecto al valor real α_0 no debe ser grande y, luego, no puede explicarse por causas aleatorias si esta desviación resulta notable, y hay que estimar que la hipótesis de que el valor del parámetro es igual a α_0 no se confirma y debe ser rechazada. Así pues, la elección del criterio de acuerdo significa la designación de la magnitud crítica de desviación $\Delta\alpha_{cr}$ elegida de tal modo que la probabilidad de sobrepasar esta magnitud sea muy pequeña. Como magnitud de la desviación crítica pueden tomarse los límites del intervalo de confianza del $100(1-q)$ por ciento. Con esto existe la probabilidad q de que la desviación observada sobrepase la crítica, lo cual significa que será rechazada injustificadamente una hipótesis correcta. La magnitud q denominada *nivel de significación estadística* debe elegirse dependiendo de las consecuencias a las cuales puede llevar el rechazo de una hipótesis correcta. Prácticamente, en la mayoría de los casos, se juzga posible tomar $q=0{,}05$, aceptando de ese modo como criterio de acuerdo un intervalo de confianza del 95%. No obstante, en los casos en que las consecuencias del rechazo incorrecto de la hipótesis pueden ser graves, se emplea $q = 0{,}01$ y a veces hasta menos.

Debe notarse que aunque el método dado permite rechazar con gran seguridad las hipótesis falsas, no puede servir como demostración de la certeza de una hipótesis. Por eso, en caso de que el valor hipotético α_0 caiga dentro de los limties del intervalo de confianza, aceptamos la hipótesis no como verdadera sino como concordante con los resultados del experimento. Para juzgar si la hipótesis es o no verdadera hay que realizar investigaciones especiales.

PROBLEMAS PARA EL CAPITULO 5

5-1. ¿Qué representan los sucesos S_4 y S_{15} en el ejemplo 5-1?

5-2. Demostrar que si el espacio de resultados del experimento contiene n elementos, en este espacio hay 2^n sucesos diferentes.

5-3. El almacén de una empresa recibe cortacircuitos de dos fábricas en la relación 1 : 3. La primera da un 10% y la segunda, un 20% de artículos defectuosos. Determinar, empleando la fórmula de la probabilidad total, la probabilidad de que resulte defectuoso un cortacircuito tomado al azar del almacén.

5-4. Construir la gráfica de la función de distribución de probabilidades para el número de tantos que cayeron al lanzar los dos dados de jugar del ejemplo 5-1.

5-5. Determinar la probabilidad de que en el ejemplo 5-5 el tiempo de espera del pasajero no exceda de 3 min.

5-6. Para saber cuántos peces hay en un lago, se pescan 1000 peces, se marcan y se sueltan de nuevo en el lago. ¿Con qué número de peces en el lago habrá la mayor probabilidad de encontrar 10 peces marcados entre los 150 peces atrapados nuevamente?

5-7. Hay que hornear 100 galletas con pasas. ¿Qué cantidad mínima de pasas hay que mezclar con la masa para que la probabilidad de que en ninguna de las galletas no haya ninguna pasa, no exceda de 0,01?

Parte segunda

OPTIMIZACIÓN DE LOS PROCESOS DE MANDO

Capítulo sexto

ESTRUCTURA Y DESCRIPCIÓN MATEMÁTICA DE LOS PROBLEMAS DE MANDO ÓPTIMO

6-1. RASGOS PRINCIPALES DEL PROCESO DE MANDO

a) Concepto de mando

Por doquier en el mundo que nos rodea (naturaleza, técnica, sociedad humana) transcurren diversos procesos cuyo carácter depende del conjunto de condiciones y factores que los acompañan. Cambiando las condiciones del transcurso de los procesos, el hombre puede influir en su carácter, cambiarlos y adaptarlos a sus objetivos. Esta intervención en la marcha natural del proceso y el cambio de la marcha natural del mismo constituye la esencia del mando. De esta manera puede decirse que *el mando representa tal organización de uno u otro proceso que garantice la consecución de determinados objetivos.*

Para una mejor comprensión de la esencia del proceso de mando examinemos el ejemplo del perro que persigue a una liebre. Para alcanzar a la liebre el perro deberá organizar sus acciones de una forma determinada, gobernándolas. Por consiguiente, el proceso de persecusión es un proceso de mando.

Al comienzo de la persecución debe preceder la aparición de la liebre, es decir, la creación de una situación durante la cual surge un objetivo determinado cuyo alcance es necesario o deseable. Pero antes de iniciar la persecución el perro debe apreciar la situación creada y compararla con sus deseos y posibilidades. La apreciación de la situación se concluye con la toma de la decisión acerca de si

conviene intentar alcanzar la liebre o no (ésta puede hallarse lejos y la persecusión es inútil, el perro puede estar agotado, etc.). Sólo después que haya tomado la decisión de perseguir, el perro comienza a organizar su movimiento fijándose como objetivo alcanzar a la liebre en el tiempo más breve o con el gasto mínimo de fuerzas.

En este ejemplo pueden distinguirse claramente las cuatro etapas características de cualquier proceso de mando: la aparición del objetivo, la estimación de la situación, la toma de decisión y la ejecución de la decisión tomada. No obstante, la etapa de aparición del objetivo precede al comienzo del proceso de mando y la excluiremos del análisis. Teniendo asimismo en cuenta que la apreciación de la situación, al mandar procesos complejos, se realiza fundándose en la información recopilada y elaborada del modo correspondiente, llegamos a las siguientes tres etapas del proceso de mando:

1) la recopilación y elaboración de la información con el objetivo de estimar la situación creada;

2) la toma de la decisión sobre las acciones más oportunas;

3) la ejecución de la decisión tomada.

A veces resulta necesaria además una cuarta etapa: el control de la ejecución de la decisión.

Los diversos tipos de problemas de mando difieren unos de otros por el método y la secuencia de ejecución de dichas operaciones.

b) Tipos de problemas de mando

Hay muchos problemas en los que los mecanismos de recopilación de la información y ejecución de la decisión tomada están elaborados con tanta precisión que puede dejar de pensarse por completo de ellos al realizar el proceso de mando. En estos problemas todo el análisis del proceso de mando, en esencia sólo se reduce al análisis de la segunda etapa. Tales problemas recibieron el nombre de problemas de toma de decisión *de una etapa* o *de un paso*.

Sin embargo, tal enfoque constituye en la mayoría de los casos una idealización y simplificación del proceso real de mando. En realidad, todas estas etapas del proceso de mando se hallan en estrecha relación mutua y la etapa de toma de decisión requiere el análisis más o menos detallado de los métodos posibles de realización de la decisión tomada. Así, para tomar la decisión de renunciar a perseguir la liebre hay que convencerse de que es inútil perseguirla y para ello es necesario analizar siquiera aproximadamente los métodos de persecución posibles.

A veces, en casos similares el proceso de mando se divide en varios pasos consecutivos, entonces la decisión tomada en cualquier paso depende de los resultados de la ejecución del paso previo. Tales procesos se denominan *de pasos múltiples* para la toma de decisión. Puede servir de ejemplo el proceso de guía de un

cohete al lanzarlo de la Tierra a la Luna. Aquí pueden distinguirse los pasos siguientes: puesta del cohete en órbita cercana a la tierra, organización del movimiento de éste en dirección a la Luna, paso del mismo a la órbita cercana a la Luna y alunizaje.

En ejemplo dado los pasos del proceso de pasos múltiples de mando se obtuvieron de modo completamente natural. Sin embargo, en muchos casos resulta una tarea sumamente difícil el fraccionamiento del complejo proceso de mando en pasos con clara diferenciación de todas las etapas de mando en cada paso. Así pues, en el proceso de persecución de la liebre nos encontramos con una situación continuamente en cambio, provocada por los esfuerzos de la liebre para escapar de la persecución. El perro deberá justipreciar continuamente esta situación y tomar ininterrumpidamente nuevas y nuevas decisiones conforme a la situación cambiable sin esperar los resultados finales de la ejecución de las decisiones anteriores. En tales problemas tropezamos con procesos de mando *continuos* y *dinámicos*.

En el análisis expuesto se ve cuán complejos y variados pueden ser los problemas de mando. Sin embargo, no estimaríamos en grado suficiente la dificultad de resolución de estos problemas si no tuviéramos en cuenta la circunstancia de que los procesos de mando transcurren comúnmente en un medio ambiente complejo. En el transcurso de los procesos de mando ejercen su influencia diversos factores externos cuyo total a menudo se denomina *estado de la naturaleza*. Para tomar una decisión correcta sobre unas u otras acciones hay que estimar los resultados de estas acciones y para ello es necesario conocer la naturaleza de la situación en la que se emprenden estas acciones.

Sin embargo, para los problemas de mando es típico el caso en que la información existente es insuficiente para la estimación precisa de la situación o está distorsionada por factores extraños. No obstante, la influencia de información no elimina el problema de tomar la decisión. La peculiaridad de los problemas consiste precisamente en que la decisión debe tomarse sin falta independientemente de si estamos en disposición de evaluar con exactitud los resultados a los que conducirá la decisión tomada.

De este modo, en el proceso de mando surge la importante tarea de tomar la decisión en las condiciones en que la información sobre la situación creada es insuficiente o distorsionada. El problema dado recibió la denominación de problema *de toma de decisión en condiciones de indeterminación*.

c) Concepto de investigación de operaciones

Anotemos una clase específica más de problemas de mando que están relacionados con las actividades de las grandes empresas industriales y pueden denominarse problemas de carácter *organizativo* y *administrativo*.

Antes de la revolución industrial podía llevar a cabo la dirección de una empresa pequeña una sola persona que hacía las compras, planeaba y orientaba el trabajo, vendía la producción, contrataba y despedía a los trabajadores. Las pequeñas dimensiones de la empresa le permitían tomar decisiones organizativas sin acudir a ningún método científico basándose en sus conocimientos, experiencia a intuición. Si algunas de las decisiones tomadas no eran las mejores, ellas no provocaban gran daño o podían ser corregidas rápidamente.

El crecimiento de las empresas industriales hizo imposible que una sola persona realizara las funciones administrativas. Surgieron los dirigentes de departamentos de producción, ventas, finanzas, personal, etc. La intensificación de la mecanización y automatización de la producción condujo al desmembramiento ulterior de las funciones administrativas. De este modo, los departamentos de producción resultaron divididos en grupos más pequeños que se ocupan de las cuestiones de explotación, reparación, control de calidad, planificación, suministros, almacenaje de la producción, etc.

Cada subdivisión especializada de una organización grande ejecuta una parte determinada del trabajo común guiándose por los objetivos generales de la empresa. No obstante, en cada subdivisión especializada surgen asimismo sus objetivos propios. Todos ellos no siempre concuerdan y a veces provocan contradicciones de unos con otros. A modo de ejemplo puede examinarse el problema de garantizar las reservas de una empresa. Una subdivisión particular puede estar interesada en aumentar considerablemente las reservas en depósito para asegurar la elaboración ininterrumpida de su produción. Pero la capacidad limitada de los lugares de depósito provoca la disminución de las reservas para otras subdivisiones. Como resultado surge un problema de tipo organizativo y administrativo: la elaboración de una estrategia tal que sea la más beneficiosa para toda la empresa en cuanto a las reservas.

Al dar solución a los problemas organizativos y administrativos similares es necesaria una precisa comprensión de los objetivos de las distintas subdivisiones y tal acuerdo entre ellos que los mismos no entren en contradicción unos con otros ni con los objetivos generales de toda la empresa. Si con esto tenemos en cuenta que la toma de decisiones no óptimas en las condiciones de una gran empresa puede provocar gran daño, se hace manifiesto que al resolver los problemas de organización y administración resulta inadmisible fundarse sólo en la experiencia personal y el sentido común. Son necesarios métodos científicos.

La joven disciplina científica que recibió el nombre de *investigación operativa* se ocupa de elaborar los métodos científicos de resolución de problemas organizativos y administrativos. En esta disciplina científica se entiende por *operación* cierta medida organizativa cuya realización persigue un objetivo determinado formulado con precisión, por ejemplo, la reglamentación de las reservas

almacenadas en un depósito. Deben presentarse las condiciones que caracterizan la situación en que se ejecutan las medidas, en particular, las necesidades de reservas y las limitaciones en locales de depósito en el ejemplo considerado. El objetivo de la investigación operativa es hallar y fundamentar científicamente aquellos métodos de realización de la medida que sean los más ventajosos en cierto sentido.

La particularidad específica de las tareas de tipo organizativo y administrativo consiste en que las consecuencias de su resolución por uno u otro método pueden reflejarse considerablemente en el trabajo de toda la empresa. Por eso la toma de la decisión definitiva siempre queda a la competencia de la persona responsable, el administrador, investido de los derechos correspondientes, y sale de los marcos de la investigación operativa. La investigación operativa sólo persigue el objetivo de poner en manos del administrador recomendaciones fundamentales para tomar la decisión.

De esta manera, la investigación operativa constituye la corriente científica cuyo objetivo es la elaboración de métodos de análisis de las acciones dirigidas a una finalidad (operaciones) y la estimación comparativa objetiva de las decisiones posibles. Aunque la investigación de operaciones constituye una corriente científica independiente que surgió en los años de la Segunda Guerra Mundial durante la resolución de los problemas de defensa antiaérea en Inglaterra (es decir, cronológicamente antes de la aparición de la cibernética), no obstante, durante la solución de algunos problemas ella aplica ampliamente los métodos de esta última.

6-2. OPTIMIZACION DEL PROCESO DE MANDO

a) Criterio de calidad de mando

En lo sucesivo vamos a considerar el problema del mando como un problema matemático. Sin embargo, a diferencia de otros muchos problemas matemáticos, éste posee la peculiaridad de que no admite una solución sino un conjunto de soluciones diferentes. Esto se debe a que en los problemas de mando, por lo general, existen muchos métodos de organización de algún proceso que conducen a alcanzar el objetivo planteado. Así, durante el proceso de persecusión de la liebre el perro puede organizar de diferente modo el carácter de su movimiento, durante el lanzamiento de un cohete a la Luna pueden elegirse distintas trayectorias para el vuelo del mismo, etc. Por eso, el problema del mando se podría plantear como el problema de buscar siquiera uno de los métodos posibles para lograr el objetivo impuesto. No obstante, tal planteo de la cuestión es, por lo común, insuficiente.

Si existe un conjunto de soluciones de cualquier problema, surge un problema adicional: elegir de este conjunto de soluciones aquella que sea la mejor desde cierto punto de vista. Pueden presentarse muchos ejemplos de estos problemas. Así pues, hay muchos métodos para pegar una caja formada de una hoja de cartón de dimensiones dadas. Puede considerarse como el problema adicional la obtención de la caja de máxima capacidad. Es posible viajar de una ciudad a otra utilizando diversos medios de transporte: aéreo, acuático, por ferrocarril, autobús y automóvil. Puede considerarse como una tarea complementaria la elección del medio de transporte más ventajoso desde el punto de vista del tiempo de viaje, el costo, la comodidad, las costumbres, etc. En los problemas de mando hay una situación análoga.

En los casos en que el objetivo del mando puede lograrse por varios métodos diferentes es posible plantear requerimientos especiales al método de mando, el grado de cumplimiento de los cuales puede servir de fundamento para dar preferencia a un método de mando sobre todos los demás.

En muchos casos la realización del proceso de mando requiere el gasto de algunos recursos: gasto de tiempo, consumo de materiales, combustible, energía eléctrica. Luego, al elegir el método hay que pensar no solamente si se consigue el objetivo planteado, sino también qué recursos hay que gastar para lograr dicho objetivo. En este caso el problema de mando consiste en elegir del conjunto de soluciones que garanticen la consecución del objetivo, la solución que requiera el gasto mínimo de recursos.

En otros casos pueden servir de fundamento para preferir un método de mando a un otro diferentes requerimientos impuestos al sistema de mando: el costo de entretenimiento, la fiabilidad, el grado de cercanía del estado obtenido del sistema al requerido, el grado de certeza de los conocimientos sobre el estado de la naturaleza, etc.

La expresión matemática que brinda una estimación cuantitativa del grado de cumplimiento de las exigencias impuestas al método de mando se denomina *criterio de calidad de mando*. El método de mando preferible u *óptimo* será aquel con el cual el criterio de calidad de mando alcanza su valor mínimo (a veces, máximo). Al elegir, por ejemplo, el régimen de vuelo puede tomarse como criterio de calidad de mando la expresión para la cantidad de combustible consumido en la unidad de espacio, o el espacio recorrido con la unidad de combustible. Al régimen más económico, o sea, óptimo, va a corresponder el valor mínimo (en el primer caso) o máximo (en el segundo caso) del criterio de calidad de mando.

Vamos a considerar como preliminar la definición de mando óptimo expuesta. Una definición más rigurosa se dará después del análisis de las limitaciones impuestas al proceso de mando.

b) Limitaciones impuestas al proceso de mando

El problema de la búsqueda del mando óptimo o del mando en general, debe considerarse inexistente, es decir, incapaz de motivar problemas algunos si no se han aplicado ningunas limitaciones al carácter del movimiento del sistema. Así, el problema de persecución de la liebre no existiría en lo absoluto si el perro pudiera salvar instantáneamente la distancia que lo separa de la liebre. Por consiguiente, al resolver un problema de mando no podemos dejar de tener en cuenta la circunstancia de que el movimiento de cualquier sistema siempre está sometido a limitaciones de diverso género.

Para representarnos más claramente las limitaciones que se manifiestan, examinemos el ejemplo concreto de conducir un automóvil. Efectuando el proceso de condución el chófer debe tomar en consideración que el automóvil tiene un motor de limitada potencia, es decir, sólo puede llevar un peso limitado con una velocidad máxima limitada. Gracias a la inercia, la velocidad del automóvil y la dirección de su movimiento pueden variar sólo con una aceleración de valor limitado. Esto significa que no es posible la detención o el cambio instantáneos de la dirección del movimiento en caso de surgir una situación peligrosa inesperada lo que a su vez limita la velocidad de movimiento. El chófer, al elegir la ruta, se ve obligado a tener en cuenta la reserva limitada de combustible en el tanque, la necesidad de reabastecerse de aquél en el camino, etc.

En el caso general hay dos tipos de limitaciones en la elección del método de mando. Las limitaciones del primer tipo son las propias leyes de la naturaleza en concordancia con las cuales ocurre el movimiento del sistema controlado. Al enunciar matemáticamente el problema del mando estas limitaciones se representan corrientemente por ecuaciones algebraicas, diferenciales o de diferencia, del objeto controlado y éstas se llaman, a menudo, ecuaciones de enlace. El segundo tipo de limitaciones es debido a la restricción de los recursos empleados durante el mando o de otras magnitudes, las cuales, a causa de las particularidades físicas de uno u otro sistema, no pueden o no deben sobrepasar de ciertos límites. Las limitaciones matemáticas de este tipo se expresan habitualmente en forma de sistemas de ecuaciones o desigualdades algebraicas las cuales se vinculan las variables que describen el estado del sistema.

c) Enunciado del problema del mando óptimo

El problema del mando puede considerarse enunciado matemáticamente si:

está enunciado el objetivo del mando expresado por medio del criterio de la calidad de mando;

están definidas las limitaciones del primer tipo que representan un sistema de ecuaciones diferenciales o de diferencias, las cuales limitan los modos posibles de movimiento del sistema;

están definidas las limitaciones del segundo tipo que representan un sistema de ecuaciones o desigualdades algebraicas que expresan la restricción de los recursos o de otras magnitudes utilizadas durante el mando.

Se denomina *mando óptimo* al método que satisface todas las limitaciones planteadas y transforme en mínimo (máximo) el criterio de la calidad de mando.

6-3. DESCRIPCION MATEMATICA DEL OBJETO CONTROLADO

a) Estructura del objeto controlado

Denominemos objeto controlado aquel sistema físico cuyos procesos mandamos. Los objetos controlados pueden ser muy variados y tener la más diversa naturaleza física. A saber:

dispositivos técnicos: automóvil, avión, cohete, torno, proceso tecnológico, etc.;

empresas industriales: departamento, taller, fábrica, rama industrial;

sistemas económicos: economía de la empresa, economía de la rama industrial, economía del estado;

sistemas biológicos;

sistemas sociales, etc.

La circunstancia de que las leyes a las cuales obedecen los procesos de mando son comunes a los objetos controlados de cualquier naturaleza física, permite estudiar la estructura general y dar una descripción matemática general del proceso de mando.

Designemos por x la variable que determina el estado del objeto controlado. A veces ésta constituye una magnitud unidimensional o escalar. Ella puede ser: el ángulo de rotación del árbol de un motor, la velocidad de un avión o cohete, la presión del vapor en la caldera de una máquina de vapor, el número de objetos en un depósito, el número de aviones con base en un aeródromo, etc.

No obstante, en la mayoría de los casos para describir el objeto controlado se requiere no una sola sino varias variables $x_1, \ldots, x_N$.

Al describir los sistemas mecánicos las magnitudes x_i representan las coordenadas o velocidades de las piezas móviles;

en los sistemas eléctricos las magnitudes x_i serán las intensidades de las corrientes o los valores de las tensiones;

en la economía ellas pueden ser las capacidades de producción o los recursos de algunas ramas industriales;

en un sistema biológico las magnitudes x_i pueden caracterizar la concentración de substancias químicas o medicinas en diversos órganos.

En todos los casos examinados el estado del objeto controlado va a describirse por una variable multidimensional, es decir, vectorial x cuyas componentes serán las magnitudes x_i:

$$x = (x_1, \ldots, x_N). \qquad (6\text{-}1)$$

En lo sucesivo denominaremos a la variable x *variable* o *vector* de estado del objeto controlado.

Las magnitudes x_i pueden variar continuamente en cierto intervalo de valores o tomar un conjunto finito de valores. En este último caso la magnitud x va a tomar también un conjunto finito de valores y vamos a designar su valor k por

$$x^{(k)} = (x_1^{(k)}, \ldots, x_N^{(k)}), \quad k = 1, \ldots, n \qquad (6\text{-}2)$$

Entonces el conjunto

$$X = \{x^{(1)}, \ldots, x^{(n)}\} \qquad (6\text{-}3)$$

va a representar el *espacio de estados posibles del objeto controlado*. A veces llamaremos al espacio X *espacio de soluciones*, subrayando así que la elección de cierta $x \in X$ constituye una solución posible del problema de mando.

Si las magnitudes x_i pueden variar continuamente, es decir, toman un conjunto infinito de valores, el espacio de estados posibles del sistema X será un conjunto infinito. Pero tampoco en este caso los valores de x_i pueden ser, por lo común, cualesquiera. Pueden aplicarse a ellos limitaciones que, como ya se indicó, tienen habitualmente la forma de sistemas de ecuaciones o desigualdades algebraicas

$$f_i(x_1, \ldots, x_N)\{\leqslant, =, \geqslant\}\, 0, \quad i = 1, \ldots, m. \qquad (6\text{-}4)$$

En cada una de las limitaciones (6-4) se conserva sólo uno de los signos $\leqslant$, $=$ ó $\geqslant$, no obstante, las distintas limitaciones pueden tener asimismo distintos signos. Las magnitudes m y N no están vinculadas entre sí, de modo que m puede ser mayor, menor o igual a N. En particular, m puede ser igual a cero, así, no se excluye el caso en que falten las limitaciones (6-4). Frecuentemente algunas o todas las variables x_i satisfacen la condición de no negatividad

$$x_i \geqslant 0, \quad i = 1, \ldots, N. \qquad (6\text{-}5)$$

La condición de no negatividad de las variables resulta muy cómoda en caso de dar solución numérica a las ecuaciones que describen el proceso de mando. Además, en muchos problemas, por ejemplo, los económicos, las magnitudes x_i no pueden ser negativas por su sentido físico (gastos, volumen de producción, cantidades de mercancías transportadas, sumas de dinero distribuidas de diversa manera, etc.). Sin embargo, los problemas en los cuales las variables pueden ser de signo arbitrario, es decir, satisfacen las limitaciones del tipo

$$x_i \geqslant a_i, \quad i = 1, \ldots, N, \qquad (6\text{-}6)$$

donde a_i son números arbitrarios que se transforman fácilmente en problemas con variables no negativas introduciendo nuevas variables:

$$y_i = x_i - a_i, \quad i = 1, \ldots, N. \tag{6-7}$$

A la par con la variable x la cual vamos a considerar en lo adelante capaz de ser medida y controlada, el estado del objeto controlado puede depender de un conjunto de factores no controlables o controlables no por completo que se determinan por un conjunto de condiciones exteriores en las que se encuentra el objeto controlado. Por ejemplo, el piloto puede regular el régimen del avión cambiando la altura y velocidad de vuelo que son los parámetros controlados en el caso dado. Sin embargo, influyen en notable grado en el consumo de combustible las condiciones atmosféricas que el piloto puede tomar en consideración sólo parcialmente pero sobre las cuales él no puede influir activamente ni siquiera preverlas con exactitud.

Vamos a denominar *estado de la naturaleza* designándolo por ϑ el total de factores exteriores no controlables que ejercen influencia sobre el proceso de mando. En muchos casos mediante cierta idealización de los fenómenos reales se logra reducir toda la variedad infinita de condiciones exteriores a un número finito de estados posibles de la naturaleza $\vartheta^{(1)}, \ldots, \vartheta^{(l)}$, o sea, introducir en el análisis el conjunto finito

$$\Theta = \{\vartheta^{(1)}, \ldots, \vartheta^{(l)}\}, \tag{6-8}$$

que se denominará en lo adelante *espacio de estados de la naturaleza*. Los elementos $\vartheta^{(i)}$ del conjunto Θ son en el caso general magnitudes multidimensionales

$$\vartheta^{(i)} = (\vartheta_1^{(i)}, \ldots, \vartheta_L^{(i)}), \; i = 1, \ldots, l. \tag{6-9}$$

La imposibilidad de controlar completamente todos los factores externos provoca que en lugar de conocer con exactitud el estado de la naturaleza ϑ en muchos casos nos vemos obligados a limitarnos solamente a conocer las probabilidades $\xi(\vartheta)$ de los diferentes estados de la naturaleza $\vartheta \in \Theta$.

Vamos a denominar *probabilidades apriorísticas* de los estados de la naturaleza a las probabilidades $\xi(\vartheta)$ obtenidas de uno u otro modo para todas las $\vartheta \in \Theta$. Naturalmente, las probabilidades apriorísticas $\xi(\vartheta)$ que representan la distribución de probabilidades en el espacio Θ deben satisfacer las condiciones (5-4).

En diversos casos es necesario considerar el estado de la naturaleza ϑ como una magnitud aleatoria continua para la cual el espacio de estados de la naturaleza Θ representa un conjunto infinito. En estos casos la distribución de las probabilidades $\xi(\vartheta)$ se convierte en la densidad de distribución de probabilidades.

Para describir la acción encaminada a una finalidad que se ejerce sobre el objeto controlado introduciremos la variable u que

vamos a denominar *acción de mando* o simplemente *mando*. Así pues, el término *mando* se utilizará en lo sucesivo en dos sentidos:

1) el mando como actividad organizativa dirigida a conseguir determinados objetivos;

2) el mando en el sentido de acción de mando, es decir, de cierta magnitud física cuyo cambio se efectúa según nuestros deseos y que influye sobre el carácter de los procesos en el objeto controlado cambiándolos en la dirección necesaria.

El sentido concreto en el cual se utiliza la palabra "mando" en los diferentes casos se aclara habitualmente por el contexto.

Durante el mando de objetos complejos es necesario utilizar, por lo común, varias acciones de mando $u_1, \ldots, u_R$, de suerte que el mando u representa en el caso general la magnitud multidimensional

$$u = (u_1, \ldots, u_R). \tag{6-10}$$

Así pues, el piloto puede cambiar el régimen de vuelo empleando el mando u que consiste en la acción sobre la cantidad de combustible consumido (u_1) y la acción sobre el timón de altura (u_2).

En los sistemas reales no pueden ser tomadas acciones de mando u_i cualesquiera sino que ellas están sometidas a limitaciones de distinto género. Designemos por U el conjunto de todos los valores de los mandos u que satisfacen las limitaciones impuestas. Entonces cualquier $u \in U$ será un *mando admisible*. En lo sucesivo van a encontrarse a menudo problemas en los que el conjunto de mandos admisibles U es el conjunto finito

$$U = \{u^{(1)}, \ldots, u^{(r)}\}. \tag{6-11}$$

Puede juzgarse sobre cómo el mando empleado u garantiza la consecución de los objetivos planteados por los valores de las variables de estado x. Sin embargo, en muchos casos esto resulta inconveniente ya que los objetivos del mando pueden expresarse por medio de las variables de estado de un mando bastante complicado. Por eso a la par con las variables de estado x conviene introducir en el análisis las *variables de salida* q que expresan en forma explícita los objetivos del mando.

A veces, en calidad de variables de salida puede utilizarse alguna de las variables de estado. Pero en el caso general esto no es así. Por ejemplo, al mandar el régimen de vuelo con la finalidad de minimizar el consumo de combustible en la ruta establecida es conveniente tomar como variable de salida el espacio recorrido por unidad de combustible, mientras que serán variables de estado la velocidad y altura del avión. Pueden ser variables de salida no sólo magnitudes físicas sino índices económicos, por ejemplo, el rendimiento, etc. En particular, el criterio de la calidad de mando también puede considerarse como variable de salida.

La variable de salida q depende en primer lugar del estado del objeto controlado x. No obstante, pueden ejercer influencia sobre ella el mando utilizado u y los factores exteriores no controlables ϑ. Por consiguiente, la ecuación para la variable de salida en el caso general tiene el aspecto:

$$q = Q(x, u, \vartheta). \qquad (6\text{-}12)$$

Puede ilustrarse la acción recíproca de todas las variables examinadas más arriba, en forma del esquema estructural del objeto controlado mostrado en la figura 6-1.

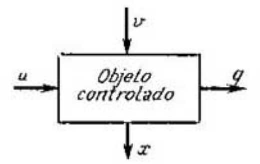


Fig. 6-1. Esquema estructural del objeto controlado

b) Ecuaciones de movimiento del objeto controlado

Bajo la acción de las señales de mando u el objeto controlado cambia su estado. El carácter de los procesos que ocurren se determina por la velocidad de cambio de la variable de estado del objeto $\dot{x} = dx/dt$ la que representa en conformidad con (6-1) la magnitud multidimensional

$$\dot{x} = (\dot{x}_1, \ldots, \dot{x}_N), \qquad (6\text{-}13)$$

donde $\dot{x}_1, \ldots, \dot{x}_N$ son las velocidades de cambio de las componentes de la variable multidimensional x.

Para los sistemas dinámicos en los cuales los procesos físicos transcurren continuamente en el tiempo, las velocidades $\dot{x}_i$ en algún momento dependen del estado del objeto controlado en el mismo momento el que a su vez se determina por los valores de la variable de estado x, el estado de la naturaleza ϑ y el mando empleado u. Esta dependencia puede escribirse en forma del sistema de ecuaciones diferenciales corrientes

$$\dot{x}_i = g_i(x, u, \vartheta), \quad x_i(0) = c_i, \quad i = 1, \ldots, N, \qquad (6\text{-}14)$$

donde las magnitudes $c_i (i = 1, \ldots, N)$ caracterizan el estado inicial del objeto controlado.

A veces resulta cómodo utilizar las notaciones vectoriales y reemplazar el sistema (6-14) por una ecuación

$$\dot{x} = g(x, u, \vartheta); \quad x(0) = c. \qquad (6\text{-}15)$$

Aqui $c = (c_1, \ldots, c_N)$ sobreentendiéndose por g la función vectorial

$$g = (g_1, \ldots, g_N). \qquad (6\text{-}16)$$

La introducción de la variable ϑ en calidad de argumento en la ecuación (6-15) en diversos casos resulta inconveniente ya que en el proceso de mando el estado de la naturaleza se mantiene por lo común invariable. Sin embargo, para distintos procesos de mando el estado de la naturaleza puede ser diferente lo que se manifiesta en el cambio de aspecto de la ecuación de movimiento. Es más cómodo señalar la dependencia del aspecto de la ecuación con relación al estado de la naturaleza sencillamente con el subíndice ϑ en la función g escribiéndose la ecuación (6-15) como

$$\dot{x} = g_{\vartheta}(u, x), \quad x(0) = c. \tag{6-17}$$

Si el estado de la naturaleza no cambia en el transcurso de todos los procesos de mando examinados entonces puede no señalarse el subíndice ϑ.

6-4. CLASIFICACION DE LOS PROBLEMAS DE MANDO OPTIMO

a) Problemas de toma de decisión de un paso

En los problemas de un paso habitualmente no se estudian los métodos de realización de la decisión tomada, es decir, no se determinan la magnitud ni el carácter de la acción de mando u, sino directamente el valor de la variable de estado del sistema x cuyo valor asegura la mejor consecución del objetivo del mando.

El problema de toma de decisión de un paso se considera planteado si están dados el espacio de estados de la naturaleza Θ con la distribución de probabilidades $\xi(\vartheta)$ para todas las $\vartheta \in \Theta$, el espacio de decisiones X y el criterio de calidad de la decisión tomada que para este caso vamos a denominar *función de objetivo.* En la literatura en lugar del término "función de objetivo" se emplea también la denominación "función de ganancias" o "función de pérdidas". La función de objetivo que expone de modo explícito los objetivos del mando puede considerarse como la magnitud de salida del objeto controlado y designarse por q. La función de objetivo constituye una magnitud escalar que depende del estado de la naturaleza ϑ y del estado del objeto controlado x pudiendo escribirse por analogía con (6-12) en la forma

$$q = Q(x, \vartheta). \tag{6-18}$$

Como vemos, el problema de toma de decisiones de un paso representa la tríada

$$G = (X, \Theta, q), \tag{6-19}$$

donde q es la función escalar definida en el producto directo de los conjuntos $X \times \Theta$. La solución de este problema consiste en hallar

una $x^* \in X$ tal que reduzca al mínimo la función q, es decir, satisfaga la condición

$$x^* = \{x \in X : Q(x, \vartheta) = \text{mín}\}. \qquad (6\text{-}20)$$

Notemos que si hay la tarea de no minimizar sino de maximizar la función q, ésta no provocará ningunas dificultades adicionales ya que si para $x = x^*$ la función $Q(x, \vartheta)$ alcanza un máximo, para la misma x la función $-Q(x, \vartheta)$ va a alcanzar un mínimo.

Existen diferentes métodos de resolución del problema de toma de decisión de un paso. La posibilidad de aplicar uno u otro método depende del modo de presentación del conjunto de las soluciones admisibles X, de la información existente sobre el estado de la naturaleza y del tipo de función de objetivo q. Daremos una característica breve de los principales de estos métodos.

El problema se llama *determinado* si no hay indeterminación en cuanto al estado de la naturaleza. En los problemas determinados el espacio de estados de la naturaleza Θ consta, por todo, de un solo elemento ϑ_0, cuya probabilidad es igual a la unidad. En este caso la función de objetivo sólo dependerá del estado del objeto controlado

$$q = Q(\vartheta^0, x) = q(x_1, \ldots, x_N). \qquad (6\text{-}21)$$

El problema determinado de un paso se denomina *problema clásico de optimización* si en ella tienen lugar limitaciones del tipo (6-4); por cierto, entre estas limitaciones no hay desigualdades, no hay condiciones de no negatividad o discresión de las variables, $m < N$ y las funciones $f_i\ (x_1, \ldots, x_N)$ y $q(x) = q(x_1, \ldots, x_N)$ son continuas que tienen derivadas parciales por lo menos de segundo orden. En este caso el problema se formula del modo siguiente. Están dadas las limitaciones del tipo

$$f_i(x_1, \ldots, x_N) = 0, \quad i = 1, \ldots, m. \qquad (6\text{-}22)$$

Hallar los valores $x_1, \ldots, x_N$ que satisfagan las ecuaciones (6-22) y que minimicen la función $q(x_1, \ldots, x_N)$.

La peculiaridad de estos problemas consiste en que ellos, al menos en principio, pueden resolverse por métodos clásicos basados en el empleo del cálculo diferencial. Efectivamente, las ecuaciones (6-22) permiten dejar de considerar m variables. Con esto la función de objetivo se reduce a la forma

$$q(x_1, \ldots, x_N) = q_1(y_1, \ldots, y_{N-m}), \qquad (6\text{-}23)$$

donde se designan por $y_1, \ldots, y_{N-m}$ las variables no excluidas. Ahora el problema consiste en encontrar los valores $y_1, \ldots, y_{N-m}$ que reducen al mínimo la función q_1 y a los cuales no se han aplicado ningunas limitaciones. Esta solución puede hallarse por el sistema de ecuaciones que se obtienen igualando a cero las derivadas parciales de la función q_1:

$$\frac{\partial q_1}{\partial y_i} = 0, \quad i = 1, \ldots, N - m. \qquad (6\text{-}24)$$

Pero, en este camino se encuentran de ordinario tales dificultades de cálculo que hacen imposible la búsqueda de otros métodos de resolución.

En los últimos años se nota un extraordinario interés hacia los métodos de resolución del problema de un paso, que han recibido el nombre de *programación matemática*. Estos métodos ofrecen la posibilidad de hallar los valores de las variables $x_1, \ldots, x_N$ que obedecen a limitaciones del tipo (6-4) tanto en forma de igualdades como de desigualdades y minimizan la función de objetivo $q(x_1, \ldots, x_N)$. Por lo general, a las variables se imponen las condiciones adicionales de que sus valores no sean negativos. Es necesario recalcar que la programación matemática no es una forma analítica sino algorítmica de resolver el problema, o sea, no proporciona una fórmula que exprese el resultado definitivo sino sólo indica el procedimiento de cálculo que conduce a la solución del problema. Por eso, los métodos de programación matemática se vuelven efectivos principalmente en caso en que se utilicen calculadoras numéricas, (digitales).

El caso más sencillo del problema de programación matemática es el de *programación lineal*. Este corresponde al caso en que los primeros miembros de las limitaciones (6-4) y la función de objetivo (6-21) representan funciones lineales de $x_1, \ldots, x_N$. En el problema de programación lineal se requiere hallar valores no negativos de las variables $x_1, \ldots, x_N$ los cuales reducen al mínimo la función de objetivo

$$q(x_1, \ldots, x_N) = \sum_{j=1}^{N} b_j x_j \tag{6-25}$$

y satisfacen el sistema de limitaciones

$$\sum_{j=1}^{N} a_{ij} x_j \leqslant 0, \quad i = 1, \ldots, m. \tag{6-26}$$

Cualquier problema de programación matemática que difiera del enunciado se denomina problema de *programación no lineal*. En los problemas de programación no lineal sea la función de objetivo (6-21), sean los primeros miembros de las limitaciones (6-4), o bien, la primera y los segundos son funciones no lineales de $x_1, \ldots, x_N$. No obstante, se considera asimismo problema de programación no lineal el problema en que la función de objetivo y las limitaciones tienen el aspecto (6-25) y (6-26), pero se presupone, por ejemplo, que las variables sean números enteros. Este último problema se denominó problema de *programación de números enteros*.

Los problemas de programación no lineal son considerablemente más complejos que los de programación lineal y en la actualidad sólo se han elaborado procedimientos de cálculo para algunos tipos de estos problemas. Más adelante examinaremos de-

talladamente los métodos de resolución de problemas de programación lineal. Los métodos de resolución de problemas de programación no lineal se estudian en otros libros al efecto.

El problema de toma de decisión de un paso se denomina *estocástico* si el espacio de estados de la naturaleza Θ consta de más de un elemento, de modo que se conoce no el estado real de la naturaleza ϑ, sino la distribución de probabilidades $\xi(\vartheta)$ en el espacio Θ.

Los problemas estocásticos que requieren hallar los valores de las variables que satisfacen las limitaciones (6-4) y minimizan la función de objetivo (6-18) se denominan problemas de programación estocástica. No obstante, en muchos casos por medio de una definición, algo distinta de la función de objetivo, los problemas de programación estocástica pueden reducrise a los problemas de programación lineal o no lineal. En efecto, como que el estado de la naturaleza ϑ es una magnitud aleatoria con distribución de probabilidades $\xi(\vartheta)$ en el espacio Θ, el valor $Q(\vartheta, x)$, dado $x = (x_1, \ldots, x_N)$ será también una magnitud aleatoria con la misma distribución de probabilidades $\xi(\vartheta)$ en el espacio Θ. Por eso en el caso dado es conveniente tomar como función de objetivo el valor medio de la función $Q(\vartheta, x)$ en el espacio Θ.

De esta manera, para los problemas estocásticos la función de objetivo según (5-64) puede ser determinada por la expresión

$$q_1(x) = \sum_{\vartheta \in \Theta} \xi(\vartheta) Q(\vartheta, x). \tag{6-27}$$

Por cuanto $q_1(x)$ es una función determinada de x, el problema de hallar las variables $x_1, \ldots, x_N$ que satisfacen las limitaciones (6-4) y minimizan la función objetiva (6-27) puede resolverse por los métodos de programación lineal o no lineal.

Un caso importante de problema estocástico de toma de decisión de un paso es el caso en que las magnitudes $x_i \times \times (i = 1, \ldots, N)$ pueden tomar solamente un conjunto finito de valores, es decir, las limitaciones para los valores de las variables están dadas en forma del espacio de decisiones X definido por la correlación (6-3). Se ocupa de los métodos para resolver estos problemas la rama de las matemáticas que ha sido denominada *teoría de las decisiones estadísticas.*

En la actualidad se presta gran atención a los problemas en los que la solución se toma no por un individuo sino por varios (por ejemplo, por dos), siendo contradictorios los intereses de éstos. Puede servir de ejemplo el problema de persecución en el que la distancia entre el persecutor y el perseguido depende de las decisiones y acciones de ambos. En este caso el persecutor está interesado en disminuir máximamente esta distancia y el perseguido, en hacerla lo mayor posible. Estos problemas recibieron la denominación de *situaciones de conflicto* y los métodos de su re-

solución se estudian en la *teoría de los juegos*. Las personas que toman las decisiones se llaman *jugadores*.

Por cuanto en una situación de conflicto cada uno de los jugadores toma decisiones independientemente de las decisiones tomadas por el otro jugador, al descubrir matemáticamente la situación de conflicto el espacio de decisiones debe considerarse como el producto directo de dos conjuntos

$$X \times Y, \tag{6-28}$$

donde $X = \{x_1, \ldots, x_n\}$ es el espacio de soluciones del primer jugador e $Y = \{y_1, \ldots, y_m\}$, el espacio de soluciones del segundo jugador.

Los elementos del espacio de decisiones $X \times Y$ van a representar pares del tipo (x, y), $x \in X$, $y \in Y$, o sea, van a determinarse por las decisiones tomadas tanto por el primer jugador, como por el segundo. Para más sencillez vamos a considerar que no existe indeterminación en el estado de la naturaleza. Entonces la función de objetivo dependerá solamente de los elementos del espacio $X \times Y$ y tendrá el aspecto

$$q = Q(x, y). \tag{6-29}$$

La contraposición de los intereses de los jugadores consiste en que el primer jugador que hace la elección de entre el conjunto X, trata de minimizar con su elección la función de objetivo, mientras que el segundo jugador, que hace la elección de entre el conjunto Y tiende a maximizarla. De esta suerte, la esencia de la situación de conflicto consiste en que cada jugador debe tomar la mejor decisión desde su punto de vista, teniendo en cuenta que su adversario hace lo mismo.

b) Problemas dinámicos de optimización de mando

Entre los diversos problemas de cibernética ocupan un lugar de importancia aquellos en que el objeto controlado se encuentra en estado de movimiento y cambio continuos bajo la influencia de diversos factores externos e internos. Los problemas de mando de tales objetos pertenecen a la clase de los *problemas dinámicos de mando*.

El objeto se denomina controlable si entre los diferentes factores que obran sobre él existen tales que disponiendo de ellos puede cambiarse la naturaleza de su movimiento. Como ya se indicó, tales acciones encaminadas a una finalidad se denominan mandos y se designan por $u(t)$.

El carácter de movimiento del objeto controlado se determina por el sistema de ecuaciones diferenciales (6-14) que es cómodo escribir abreviadamente al estilo vectorial en forma de una ecuación diferencial (6-17). El mando $u(t)$ forma parte del mando (6-17), de modo que esta ecuación determina no sólo el movi-

miento concreto del objeto sino sus posibilidadas técnicas que pueden ser realizadas mediante el empleo de uno u otro mando del espacio de mandos admisibles U.

Como antes, puede evaluarse en qué grado se logran con uno u otro método de mando los objetivos planteados introduciendo la función de objetivo del tipo (6-12), que en el caso dado es conveniente escribir en la forma

$$q = Q_{\vartheta}[x(t),\ u(t)]. \qquad (6\text{-}30)$$

Así pues, si $u(t)$ es el consumo instantáneo de combustible y $x(t)$, la velocidad instantánea de un avión, desde el punto de vista del consumo de combustible, la calidad de mando en cualquier momento puede ser caracterizada por la magnitud $q(t) = u(t)/x(t)$ (consumo instantáneo de combustible en la unidad de espacio), que, como es natural, va a depender del estado de la naturaleza ϑ, o sea, de los factores externos que determinan las condiciones de vuelo.

La función de objetivo de tipo (6-30) se aplica raramente ya que la misma sólo ofrece la estimación de los valores instantáneos del proceso mandado cuando en la mayoría de los problemas resulta necesario estimar procesos en el objeto de mando durante todo el tiempo de mando de 0 a T.

En muchos casos se logra seleccionar una función de objetivo tal que la estimación del proceso en el objeto controlado pueda efectuarse integrando la función de objetivo en todo el tiempo de mando, o sea, tomando como criterio de calidad de mando la funcional

$$J(u) = \int_0^T Q_{\vartheta}[x(t),\ u(t)]\,dt. \qquad (6\text{-}31)$$

Así pues, si la función de objetivo tiene el sentido físico de las pérdidas, la expresión (6-31) determina la suma de pérdidas durante todo el proceso de mando.

A veces, en calidad de objetivo de mando se logra presentar la marcha deseada del proceso $z(t)$. Entonces, en calidad de función de objetivo puede tomarse el cuadrado o la magnitud absoluta de desviación del proceso $x(t)$ respecto al deseado

$$q = [x(t) - z(t)]^2, \quad q = |x(t) - z(t)|. \qquad (6\text{-}32)$$

En estos casos el criterio de calidad de mando (6-31) va a determinar el error total cuadrático o el absoluto.

En los problemas dinámicos de mando a la par con las limitaciones del tipo (6-11), las cuales determinan el espacio de mandos admisibles U, nos vemos obligados a operar con limitaciones integrales del tipo

$$\int_0^T H_{\vartheta}[x(t),\ u(t)]\,dt \leqslant K = \text{const}. \qquad (6\text{-}33)$$

Por ejemplo, muy a menudo tropezamos con la necesidad de acotar los límites de cambio del valor instantáneo de cierto parámetro $a(x, u)$ durante el proceso de mando. Designemos por a_0 aquel valor del parámetro a que es indeseable que se sobrepase. Si la función subintegral $H(x, u)$ se determina por la correlación

$$H(x, u) = \begin{cases} 0 & \text{para } a(x, u) \leqslant a_0; \\ [a(x, u) - a_0]^2 & \text{para } a(x, u) > a_0, \end{cases} \tag{6-34}$$

la limitación integral (6-33) va a expresar el requerimiento de que el valor instantáneo del parámetro a pueda sobrepasar de a_0 únicamente durante un corto tiempo y en un valor insignificante. Esta condición va a cumplirse tanto más rígidamente, cuanto menos sea K. Así pues, para $K = 0$, la limitación (6-33) no admitirá del todo que a exceda de a_0.

Las limitaciones del tipo (6-33) surgen asimismo cuando tenemos que operar con recursos limitados: puede estar limitada, si se trata de trayectorias, la cantidad de energía, combustible de que se dispone, etc.

Las correlaciones expuestas permiten dar la siguiente definición del mando óptimo en los sistemas dinámicos. Se denomina *óptimo* el mando $u^*(t)$ elegido del espacio de mandos admisibles U, que para el objeto descrito por la ecuación (6-17) minimiza el criterio de calidad (6-31) con las limitaciones presentadas a los recursos a utilizar (6-33).

Los problemas dinámicos de mando, al igual que los de un paso, pueden ser *determinados*, si el espacio de estados de la naturaleza Θ consta de un solo elemento $\vartheta = \vartheta_0$ y *estocásticos*, si dicho espacio Θ contiene más de un elemento y se ha estipulado la distribución apriorística de las probabilidades $\xi(\vartheta)$ en el espacio Θ.

Entre los problemas estocásticos ocupan un importante lugar los de mando *adaptable* que se utiliza en los casos en que los datos apriorísticos sobre el estado de la naturaleza resultan insuficientes para llevar a cabo un mando eficaz, o falta una descripción matemática suficientemente exacta del propio objeto controlado. El mando adaptable tiene como finalidad precisar los datos acerca del estado de la naturaleza o las propiedades del objeto controlado directamente en el proceso de mando del objeto probando los distintos métodos de mando y buscando aquel de ellos que resulte más eficaz en las condiciones concretas dadas.

c) Mando del estado final

En toda una serie de casos la naturaleza del movimiento del objeto durante el proceso de mando no presenta un interés substancial y sólo es de importancia el estado que tomará el objeto en el momento en que termina el proceso de mando. Pueden servir de

ejemplos de problemas similares el aterrizaje "a ciegas" de un avión que plantea rigurosas exigencias a la velocidad y la posición del avión en el momento en que toca tierra, el transporte de carga para un plazo dado a un punto de destinación dado, el alcance a finales del año de una productividad del trabajo, etc. Estos problemas recibieron la denominación de problemas *de mando del estado final*.

Designemos por $x(T)$ el estado del objeto en el momento final. En el caso dado la función de objetivo va a tener la forma

$$q = Q_{\vartheta}[x(T)]. \tag{6-35}$$

Por cuanto $x(T)$ depende de la naturaleza del mando empleado $u(t)$, el valor q también dependerá del mismo. Por eso el problema de la elección del mando óptimo puede enunciarse para este caso del modo siguiente: elegir tal mando $u^*(t)$ del espacio de mandos admisibles U que para el objeto descrito por la ecuación diferencial (6-17) minimice la función de objetivo (6-35) con las limitaciones (6-33) de los recursos a utilizar.

d) Juegos diferenciales

La teoría de los juegos diferenciales constituye la extensión de la teoría de los juegos en los problemas de un paso al caso de los problemas dinámicos de mando. En los juegos diferenciales, como en los habituales, toman parte dos personas que se llaman jugadores y cuyos intereses están contrapuestos. Con esto, cada uno de los jugadores puede dominar hasta cierto punto la situación influyendo sobre el objeto con su mando propio. Esto significa que en los juegos diferenciales el espacio de mandos admisibles representa el producto directo de los conjuntos

$$U \times V, \tag{6-36}$$

donde U es el espacio de mandos admisibles del primer jugador con los elementos $u(t)$ y V es el espacio de mandos admisibles del segundo jugador con los elementos $v(t)$. De esta manera, los elementos del espacio de mandos admisibles van a representar pares del tipo (u, v), $u \in U$, $v \in V$.

Si no existe indeterminación en el estado de la naturaleza la ecuación diferencial, que describe el movimiento del objeto controlado, se obtiene de (6-17) sustituyendo u por el par (u, v), es decir, tendrá la forma:

$$\dot{x} = g(x, u, v), \quad x(0) = c. \tag{6-37}$$

Los juegos diferenciales son de ordinario problemas para el mando del valor final. Por eso, la función de objetivo depende del estado del objeto en el momento final T que a su vez se determina por las ecuaciones utilizadas, o sea, puede escribirse en la forma:

$$q = Q[x(T)], \quad x(T) = f[u(t), v(t)]. \tag{6-38}$$

La contraposición de los objetivos de los jugadores consiste en lo siguiente: mientras el primer jugador tiende a minimizar la función de objetivo eligiendo el mando u del espacio U, los intereses del segundo requieren tal elección del mando v del espacio V, que maximice la función de objetivo. La esencia de la teoría de los juegos diferenciales es la búsqueda por cada uno de los jugadores, en esta situación tan contradictoria, de los métodos de determinación de su mando óptimo.

Son ejemplos de situaciones características para los juegos diferenciales, las batallas, los combates aéreos, el fútbol, la persecución de un buque por un torpedo, la protección de objetos contra asaltos, etc.

6-5. PROCESOS DE MANDO DE PASOS MULTIPLES

a) Comportamiento de un sistema dinámico como función del estado inicial

La búsqueda del mando óptimo en los sistemas dinámicos se facilita considerablemente en muchos casos si se logra fraccionar el proceso de mando de modo natural o artificial en pasos o etapas separados.

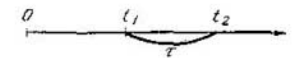


Fig. 6-2. Fraccionamiento del intervalo en dos pasos

Para efectuar el análisis en forma general, vamos a considerar que el estado del objeto se describe con la variable multidimensional

$$x = (x^{(1)}, \ldots, x^{(n)}). \tag{6-39}$$

Suponiendo que el proceso no es controlado y que no hay indeterminación en el estado de la naturaleza, la ecuación diferencial que define el movimiento del objeto con (6-17) se escribirá por analogía en la forma

$$\dot{x} = g(x), \;\; x(0) = c. \tag{6-40}$$

La solución de esta ecuación se escribe corrientemente como $x = x(t)$ con lo que se recalca que la solución depende del tiempo. Sin embargo, no es menos importante que la solución de la ecuación (6-40) depende del estado inicial c. Por eso es más rigurosa la forma de notación que muestra la dependencia explícita de la solución x tanto del tiempo como del estado inicial:

$$x = x(c, t) = x[x(0), t]. \tag{6-41}$$

Esta forma de notación permite considerar el estado del sistema en el momento de tiempo arbitrario t como cierta transformación del estado inicial $x(0) = c$ en el intervalo t.

Examinemos el movimiento del objeto en el intervalo de 0 a t_2 que dividiremos con el punto intermedio t_1 en dos intervalos de duracion t_1 y $\tau = t_2 - t_1$ como se muestra en la figura 6-2. Examinemos tres estados del objeto controlado:

estado inicial $x(0) = c$;

estado $x(c, t_1)$ en el momento intermedio t_1;

estado $x(c, t_2)$ en el momento final t_2.

Puede emplearse la descripción del último estado de doble manera. Dicho estado puede considerarse como la transformación del estado inicial $x(0) = c$ en el intervalo $t_2 = t_1 + \tau$

$$x(c, t_2) = x(c, t_1 + \tau), \qquad (6\text{-}42)$$

o como la transformación del estado $x(c, t_1)$ en el intervalo τ

$$x(c, t_2) = x[x(c, t_1), \tau] \qquad (6\text{-}43)$$

Como que ambas expresiones describen un mismo estado, entonces, igualándolas, obtenemos la correlación

$$x(c, t_1 + \tau) = x[x(c, t_1), \tau]. \qquad (6\text{-}44)$$

b) Representación de un proceso dinámico en forma de sucesión de transformaciones

Supongamos que el proceso dinámico $x(c, t)$ en el intervalo de 0 a t' puede representarse de modo natural o artificial como pasos múltiples y hallemos el método apropiado de descripción de dicho proceso. Para obtener el proceso de pasos múltiples el intervalo



Fig. 6-3. Fraccionamiento del intervalo en n pasos

de 0 a t' debe ser fraccionado en n pasos consecutivos cuya duración tomaremos igual a $\tau_1, \tau_2, \ldots, \tau_n$ como se muestra en la figura 6-3. Designemos por t_k ($k = 0, \ldots, n$) los momentos de terminación del k-ésimo paso de modo que $t_{k+1} = t_k + \tau_{k+1}$ y por x_k el estado del objeto en el momento t_k:

$$x_k = x(c, t_k). \qquad (6\text{-}45)$$

El estado

$$x_{k+1} = y(c, t_{k+1}) = x(c, t_k + \tau_{k+1}). \qquad (6\text{-}46)$$

Esta expresión puede representarse en concordancia con (6-44) y (6-45) en la forma

$$x_{k+1} = x[x(c, t_k), \tau_{k+1}] = x(x_k, \tau_{k+1}). \qquad (6\text{-}47)$$

La correlación (6-47) representa el estado del objeto x_{k+1} como resultado de la transformación del estado x_k en el paso $(k + 1)$.

Introduzcamos en el análisis el operador T que va a designar la transformación del estado del proceso durante un paso:

$$T(x_k) = x(x_k, \tau_{k+1}), \quad k = 0, \ldots, n-1. \tag{6-48}$$

Entonces la correlación (6-47) se escribirá en la forma

$$x_{k+1} = T(x_k). \tag{6-49}$$

Suponiendo que $k = 0, 1, \ldots, n-1$ podemos escribir todo el proceso dinámico en forma de la sucesión de transformaciones

$$x_0 = c, \ x_1 = T(x_0), \ \ldots, \ x_n = T(x_{n-1}). \tag{6-50}$$

c) Proceso de mando de pasos múltiples

El proceso dinámico descrito por la transformación (6-49) no es controlado. Para obtener un proceso controlado de pasos múltiples hay que tener la posibilidad de realizar en cada paso no una transformación $T(x_k)$ sino una del conjunto de transformaciones $T_1(x_k), \ldots, T_r(x_k)$.

Es cómodo considerar que el tipo concreto de transformación dependerá del parámetro u_k que en el k-ésimo paso puede tomar

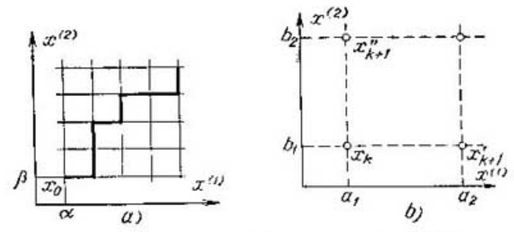


Fig. 6-4. Proceso de pasos múltiples con variable bidimensional

uno de los valores del conjunto U_k. Vamos a denominar el parámetro u_k mando y el conjunto U_k espacio de mandos admisibles en el k-ésimo paso. Ahora puede escribirse la transformación realizada en el k-ésimo paso en la forma

$$x_{k+1} = T(x_k, u_k), \ u_k \in U_k. \tag{6-51}$$

Si en la correlación (6-51) hacemos consecutivamente $k = 0, 1, \ldots, n-1$ y tomamos en cuenta el estado inicial x_0, obtendremos la descripción de todo el proceso de mando de pasos múltiples:

$$x_{k+1} = T(x_k, u_k), \ u_k \in U_k;$$
$$k = 0, 1, \ldots, n-1; \ x_0 = x(0) = c. \tag{6-52}$$

La correlación (6-52) denominada ecuación de diferencias del objeto controlado es análoga a la ecuación diferencial (6-17) que proporciona la descripción del proceso dinámico continuo.

Ejemplo 6-1. Supongamos que la variable de estado es una magnitud bidimensional $x = (x^{(1)}, x^{(2)})$ que puede tomar los valores determinados geométricamente por los nudos de la retícula representada en la figura 6-4, *a*. La transición de un nudo de retícula al siguiente se realiza empleando en cada paso uno de los dos mandos posibles: $u_k = 0$, movimiento en sentido horizontal y $u_k = 1$, movimiento en sentido vertical. Por consiguiente, el espacio de mandos admisibles, igual para cualquier paso, es

$$U_k = \{0,\ 1\},\ k = 0,\ 1,\ \ldots,\ n-1.$$

Examinemos uno de los cuadrados de la retícula dada que se muestra en la figura 6-4, *b* en cuyo ángulo inferior izquierdo el sistema se vio después del k-ésimo paso, de modo que $x_k = (a_1, b_1)$. El valor $x_{k+1} = T(x_k, u_k)$ depende de la ecuación aplicada. Como se ve, en la figura 6-4, *b* tienen lugar las siguientes correlaciones:

$$T[(a_1,\ b_1),\ 0] = (a_2,\ b_1) = x'_{k+1};$$

$$T[(a_1,\ b_1),\ 1] = (a_1,\ b_2) = x''_{k+1}.$$

La trayectoria concreta de movimiento del sistema puede describirse indicando el estado inincial x_0 y la succeción de ecuaciones empleadas. Así, la trayectoria marcada por la línea gruesa en la figura 6-4, *a*, se obtiene utilizando el mando $u = (01101001)$ con las condiciones iniciales $x_0 = (\alpha, \beta)$.

d) Criterio de calidad del mando en el proceso de pasos múltiples

Vamos a considerar que la calidad del mando se determina por el valor de la función de objetivo q, cuyo valor numérico puede considerarse como pérdidas que tenemos empleando uno u otro mando. Las pérdidas en un paso van a depender del estado del proceso al comienzo del paso y del mando empleado en este paso, es decir,

$$q_k = Q(x_k,\ u_k),\ u_k \in U_k. \qquad (6\text{-}53)$$

Como criterio de calidad del mando pueden tomarse las pérdidas completas durante todos los n pasos del proceso y representar el criterio de calidad de mando del proceso de n-ésimo paso en la forma

$$J_n(u) = \sum_{k=0}^{n-1} Q(x_k,\ u_k). \qquad (6\text{-}54)$$

Aquí por u se designa la sucesión de los mandos, o sea, un conjunto ordenado de tipo

$$u = (u_0,\ \ldots,\ u_{n-1}),\quad u_0 \in U_0,\ \ldots,\ u_{n-1} \in U_{n-1}. \qquad (6\text{-}55)$$

Si se estudia el proceso de pasos múltiples de mando del valor final, las pérdidas van a depender solamente del estado del objeto

controlado al final del proceso cuyo estado depende, a su vez, del estado inicial y de los mandos empleados en cada paso:

$$q = q(x_n) = Q(x_0,\ u), \qquad (6\text{-}56)$$

donde u se determina por la correlación (6-55). Vamos a denominar *función de objetivo* del proceso de mando de pasos múltiples a la magnitud escalar determinada por la expresión (6-56).

El problema de la búsqueda del mando óptimo en el proceso de pasos múltiples puede formularse del modo siguiente. Para el sistema dinámico, cuyos procesos se describen con la ecuación de diferencias (6-52), hallar tal sucesión de ecuaciones $u_0, u_1, \ldots, u_{n-1}$ que obedezcan a las limitaciones del tipo $u_k \in U_k$, $k = 0, 1, \ldots, n-1$, la cual minimiza el criterio de calidad de mando (6-54) o la función de objetivo (6-56).

Ya que los procesos de mando de pasos múltiples son un caso particular de procesos dinámicos de mando, entre ellos pueden encontrarse problemas de todos aquellos tipos que se estudiaron en los problemas de mando dinámico. Así pues, si el espacio de estados de la naturaleza Θ consta de un solo elemento ϑ_0, el problema de mando de pasos múltiples se denomina *determinado*. En caso contrario éste pertenece a la clase de los problemas *estocásticos*.

Subrayemos que no siempre es posible incluir unívocamente uno u otro problema en la clase de los problemas de pasos múltiples o de un paso. Así pues, si en un problema de pasos múltiples la sucesión de valores de la variable que se toma en los pasos aislados se considera como una variable multidimensional, el problema de pasos múltiples se convierte en problema de un paso.

6-6. PROBLEMA DETERMINADO DE OPTIMIZACION DE UN PASO CON VARIABLE UNIDIMENSIONAL DE ESTADO

a) Enunciado del problema

Entre los variados problemas de optimización del mando el de un paso con variable unidimensional ocupa un lugar un tanto particular. Esto es debido, en primera, a que en la práctica tales problemas se encuentran con mucha frecuencia y en segunda, muchos problemas de pasos múltiples y de un paso con variable multidimensional incluyen en sí en calidad de etapas separadas de la resolución un problema de un paso con variable unidimensional. Aunque el problema de optimización de un paso con variable unidimensional puede ser resuelto en la mayoría de los casos por los medios del análisis matemático que se estudia en el curso de matemáticas superiores, no obstante, el esclarecimiento de algunas propiedades de la resolución y el empleo de algunos procedimien-

tos de cálculo permiten en muchos casos simplificar notablemente el proceso de resolución.

El problema determinado de optimización de un paso con variable unidimensional puede formularse del modo siguiente. Sea R el conjunto de números reales; $X \subseteq R$, el conjunto de los valores admisibles de la variable unidimensional; y x, $q(x)$, una función de x estipulada en el conjunto X. Se requiere hallar tal valor de $x \in X$ que se designará en lo adelante por x^* que convierte $q(x)$ en el mínimo o máximo, es decir, satisface la condición

$$x^* = \{x \in X : q(x) = \text{mín}\,(\text{máx})\}. \qquad (6\text{-}57)$$

Los métodos de resolución de este problema dependen del carácter del conjunto X de valores admisibles de la variable que vamos a suponer presentado en un intervalo real finito. Estudiemos algunos casos particulares.

b) Caso del conjunto finito de soluciones admisibles

Analicemos el caso en que el conjunto de soluciones admisibles es finito

$$X = \{x_1, \ldots, x_N\}. \qquad (6\text{-}58)$$

Con esto la función de objetivo $q(x)$ se representará por el conjunto de números reales Q que pueden considerarse como el reflejo del conjunto X en el conjunto de los números reales R:

$$Q:\ X \to R. \qquad (6\text{-}59)$$

Los elementos del conjunto Q, es decir, los valores $q(x)$ para todas $x \in X$ pueden hallarse por medio de cálculos.

El procedimiento general para obtener el valor óptimo x^* consiste en comparar entre sí por pares todos los elementos del conjunto Q y buscar entre ellos el elemento menor o mayor. Tal procedimiento no es complicado si el conjunto X, y por ende, el conjunto Q tienen un pequeño número de elementos. Si el número de elementos del conjunto X es grande, el procedimiento dado requerirá un considerable gasto de tiempo lo que va asociado, en primer lugar, a la necesidad de calcular muchas veces el valor $q(x)$.

El único modo de simplificar este procedimiento sólo puede ser el intentar a disminuir el número de elementos del conjunto X eliminando aquellos elementos con respecto a los cuales podemos estar convencidos [no fundándose en el cálculo de los valores $q(x)$ sino basándose en indicios secundarios cualesquiera], de que entre ellos no puede estar contenido el valor x^*. Más adelante se pondrán ejemplos del empleo de este procedimiento que, claro está, siempre debe tenerse presente al resolver problemas con variables multidimensionales.

c) Caso del conjunto infinito limitado de soluciones admisibles

El caso que se encuentra con más frecuencia es aquel en que el conjunto X de valores admisibles de la variable representa el intervalo real continuo limitado por los números a y b:

$$X = \{x \in R : a \leqslant x \leqslant b\} = [a, b]. \tag{6-60}$$

Vamos a suponer que la función $q(x)$ es continua y diferenciable en cada punto del intervalo $[a, b]$ con la excepción posible de un número finito de puntos. En la figura 6-5 se muestra un ejemplo de función que satisface este requisito.

Designemos por x^*, como antes, el valor $x \in X$ con el cual se alcanza el mínimo o el máximo de la función $q(x)$. Como fundamento para la determinación de x^* ponemos el conocido postulado del curso de análisis matemático consistente en que para una función continua y diferenciable sólo puede tener lugar máximo o mínimo en aquellos puntos donde la derivada es cero. Estos puntos pueden hallarse por la ecuación

$$q'(x) = 0. \tag{6-61}$$

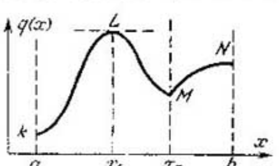


Fig. 6-5. Ejemplo de función de objetivo

No obstante, en el intervalo $[a, b]$ pueden encontrarse puntos en los cuales la derivada no existe, como por ejemplo el punto M en la figura 6-5, y también los extremos del intervalo a y b. En estos puntos que no obedecen a la condición (6-61) puede hallarse asimismo un máximo o mínimo de la función $q(x)$.

Para formular la regla general introduzcamos las siguientes notaciones. Designemos por

$$S_1 = \{x \in X : q'(x) = 0\} \tag{6-62}$$

el conjunto de puntos del intervalo $[a, b]$, en los que la derivada se reduce a cero;

$$S_2 = \{x \in X : q'(x) \text{ no existe}\} \tag{6-63}$$

el conjunto de puntos del intervalo $[a, b]$, en los cuales no existe la derivada;

$$S_3 = \{a, b\} \tag{6-64}$$

el conjunto de puntos consistentes en los extremos del intervalo $[a, b]$.

El valor x^* que minimiza o maximiza $q(x)$ se encontrará, sin falta, entre los puntos del conjunto S_1, entre los puntos del conjunto S_2 o entre los puntos del conjunto S_3. En otras palabras, el valor $x = x^*$ debe buscarse entre los puntos $\tilde{x}$ del conjunto S de-

terminado por la condición

$$S = S_1 \cup S_2 \cup S_3 \qquad (6\text{-}65)$$

que es un conjunto finito. Esta regla reduce el problema de buscar el valor óptimo x^*, al análisis del caso anterior del conjunto finito de soluciones admisibles.

Ejemplo 6-2. Para la función dada en la figura 6-5 tenemos $S_1 = \{x_1\}$, $S_2 = \{x_2\}$, $S_3 = \{a, b\}$.

Luego, $S = \{a, x_1, x_2, b\}$.

Ejemplo 6-3. Sean $X = [0, 2]$, $q(x) = \left| x - \frac{1}{2} \right|$. Como según la definición de la magnitud absoluta de la y real

$$|y| = \begin{cases} -y \text{ para } y < 0, \\ y \text{ para } y \geqslant 0, \end{cases}$$

$q(x)$ puede representarse en la forma

$$q(x) = \begin{cases} \frac{1}{2} - x \text{ para } x < \frac{1}{2}; \\ x - \frac{1}{2} \text{ para } x \geqslant \frac{1}{2}. \end{cases}$$

Diferenciando por x, obtenemos:

$$q'(x) = \begin{cases} -1 \text{ para } x < \frac{1}{2}; \\ +1 \text{ para } x > \frac{1}{2}. \end{cases}$$

En el punto $x = 1/2$ la derivada no existe, ya que los valores $q'(x)$ en las cercanías de este punto son diferentes a la izquierda y a la derecha. Como vemos, en el intervalo $[0, 2]$ no hay valores de x para los cuales $q'(x) = 0$. De esta suerte, $S_1 = \varnothing$, $S_2 = \{1/2\}$, $S_3 = \{0, 2\}$ y $S = \{0, 1/2, 2\}$.

Calculando los valores $q(x)$ para todas las $x \in S$, obtenemos $\min\limits_{x \in [0, 2]} q(x) = 0$ para $x = 1/2$, $\max\limits_{x \in [0, 2]} q(x) = 3/2$ para $x = 2$.

Hay que recalcar que los valores $\bar{x} \in S$ determinan los mínimos y máximos relativos o locales, es decir, aquellos puntos $\bar{x}$ del conjunto X para los que

$$q(\bar{x}) < q(x) \quad \text{o} \quad q(\bar{x}) > q(x)$$

para todas las x suficientemente próximas a $\bar{x}$. Es cierto que entre los puntos $\bar{x} \in S$ pueden hallarse también algunos en los que no hay máximo ni mínimo relativos, por ejemplo, los puntos de inflexión. El cálculo de los valores $q(x)$ en los puntos $x = \bar{x} \in S$ tiene el objetivo de separar el mínimo o máximo absolutos de la función $q(x)$. Con esto, el procedimiento de resolución descrito anteriormente puede resultar bastante complejo si el número de elementos del conjunto S es grande y la función $q(x)$, complicada para los cálculos.

La simplificación de la resolución del problema consiste en disminuir el número de elementos del conjunto S lo que puede lograrse cambiando el cálculo de la función $q(x)$ por el cálculo de las derivadas si éstas tienen una expresión más sencilla que la propia función. Con esto pueden aplicarse los métodos de eliminación, conocidos del curso de matemáticas superiores, de algunos elementos del conjunto S basados en la comparación de los signos de las derivadas o en el cálculo de las derivadas de órdenes superiores.

Ejemplo 6-4. Hallar el valor de $x = x^*$ que maximiza la función $q(x) = e^{-x^2}$ en el intervalo $X = [-1, +1]$

De la condición $q'(x) = -2xe^{-x^2} = 0$ hallamos que $S_1 = \{0\}$. Por cuanto la función $q(x)$ es continua, $S_2 = \varnothing$ y $S_3 = \{-1, +1\}$. De esta manera, $S = \{-1, 0, +1\}$.

Tratemos de disminuir el número de elementos del conjunto S. Por cuanto

$$q'(x) = \begin{cases} \geqslant 0 \text{ para } x < 0; \\ < 0 \text{ para } x > 0, \end{cases}$$

en los puntos del conjunto S_3 no puede haber máximo y para el punto $x = 0$ se satisfacen las condiciones de máximo. Por tanto, $S' = S_1 = \{0\}$, $x^* = 0$, $q(x^*) = e^0 = 1$.

En el siguiente ejemplo examinaremos las dificultades que surgen al resolver problemas con valores numéricos enteros de las variables.

Ejemplo 6-5. Hallar el valor de $i = i^*$ que maximiza la función $q(i) = 100i - 3i^2$ en el conjunto $X = \{i \in R : i, \text{ entero}; 0 \leqslant i \leqslant 100\}$.

Aquí el dominio de valores admisibles de la variable representa un conjunto finito. Sin embargo, este último contiene un número tan grande de elementos, que resulta difícil comparar directamente los valores $q(i)$ para todos los elementos del conjunto X. En estos problemas siempre hay que tratar de extender el dominio de valores admisibles de la variable de suerte que ella represente un intervalo continuo. Por ejemplo, puede sustituirse X por $X_1 = \{x \in R : 0 \leqslant x \leqslant 100\}$ de modo que $X \subset X_1$. Si ahora, en lugar de $q(i)$ se construye una nueva función $q_1(x)$ tal que para $x = iq_1(x) = q(i)$, puede reemplazarse el problema de hallar la $x^* \in X$ que maximice $q(i)$ por el de encontrar $x^* \in X_1$ que maximice $q_1(x)$.

Hagamos $q_1(x) = 100x - 3x^2$ en el ejemplo considerado. Entonces $S_1 = \{16{,}7\}$, $S_2 = \varnothing$ y $S_3 = \{0{,}; 100\}$. De esta forma, $S = \{0; 16{,}7; 100\}$ y $x^* = 16{,}7$. Sin embargo, el valor de x^* obtenido no es elemento del conjunto X. Por eso hay que hallar el valor óptimo de i^* entre los elementos del conjunto X adyacentes a x^*. Así pues, obtenemos $i^* \in \{16; 17\}$ de donde hallamos que $i^* = 17$, $q(i^*) = 833$.

d) Aplicación de las fórmulas de interpolación

En diversos casos, como por ejemplo, al optimizar algunas etapas de un proceso de pasos múltiples la función $q(x)$ no se presenta a modo de fórmula sino que se indica únicamente el procedimiento de obtención de los valores $q(x)$. En estos problemas resulta imposible emplear los métodos analíticos descritos en la sección precedente y para juzgar sobre la naturaleza de la función

$q(x)$ hay que encontrar todos los valores tomados por esta función en el intervalo de los valores admisibles de la variable $X = [a, b]$. Por cuanto, es evidentemente imposible exponer en forma tabular todos los valores de la función $q(x)$, se requiere calcular $q(x)$ para cierto subconjunto finito $\mathring{X}$ del conjunto X y después utilizar los resultados obtenidos para darse una idea de cómo se comporta esta función en todo el intervalo $[a, b]$. Este procedimiento se denomina interpolación de la función $q(x)$. Sin detenernos en los problemas generales de la teoría de la interpolación, examinemos sólo una de las fórmulas de interpolación más difundidas denominada fórmula de Newton.

Supongamos que en el intervalo $[a, b]$ se han tomado $n + 1$ valores discretos equidistantes de x que constituyen el conjunto $\mathring{X} = \{x_0 = a, x_1, \ldots, x_n = b\}$ siendo $x_i = x_0 + i\delta$, $i = 0, 1, \ldots, n$, donde δ es el intervalo entre los valores discretos adyacentes.

Vamos a considerar que se han calculado los valores $q(x)$ para $x \in \mathring{X}$ que son iguales a $y_i = q(x_i)$, $i = 0, 1, \ldots, n$. La interpolación consiste en que se elige el polinomio $p(x)$ de un grado no superior a n que toma en los puntos $x = x_i$ los valores $p(x_i) = y_i$. Los valores de este polinomio para cualquier $x \in [a, b]$ se toman como valores de $q(x)$.

Escribamos el polinomio $p(x)$ en la forma

$$p(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1) + \\ + a_3(x - x_0)(x - x_1)(x - x_2) + \ldots \\ + a_n(x - x_0)(x - x_1)\ldots(x - x_{n-1}) \quad (6\text{-}66)$$

Determinemos los coeficientes $a_0, a_1, \ldots, a_n$ de suerte que

$$p(x_0) = y_0, \quad p(x_1) = y_1, \ldots, p(x_n) = y_n.$$

Sustituyendo en (6-66) en lugar de x consecutivamente los valores $x_0, x_1, \ldots, x_n$, obtenemos que

$$y_0 = a_0 \quad \text{o} \quad a_0 = y_0;$$

$$y_1 = a_0 + a_1(x_1 - x_0) = y_0 + a_1\delta \quad \text{o} \quad a_1 = \frac{y_1 - y_0}{\delta};$$

$$y_2 = a_0 + a_1(x_2 - x_0) + a_2(x_2 - x_0)(x_2 - x_1) = \\ = y_0 + 2(y_1 - y_0) + 2\delta^2 a_2 \quad \text{o} \quad a_2 = \frac{y_2 - 2y_1 + y_0}{2\delta^2}.$$

Este proceso de cálculo consecutivo de coeficientes puede continuarse más adelante.

Para la más cómoda notación de los coeficientes es conveniente introducir designaciones especiales para las diferencias de valores de las funciones. Se denominan primeras diferencias las magnitudes

$$\Delta y_k = y_{k+1} - y_k, \quad k = 0, \quad 1, \ldots, n - 1.$$

Las diferencias de las primeras diferencias se denominan diferencias secundarias y se designan por

$$\Delta^2 y_k = \Delta y_{k+1} - \Delta y_k = y_{k+2} - 2y_{k+1} + y_k.$$

Análogamente pueden introducirse las diferencias tercera, cuarta, etc. Por ejemplo,

$$\Delta^3 y_k = \Delta^2 y_{k+1} - \Delta^2 y_k = y_{k+3} - 3y_{k+2} + 3y_{k+1} - y_k.$$

Teniendo en cuenta estas designaciones las fórmulas para los coeficientes toman la forma:

$$a_0 = y_0; \quad a_1 = \frac{\Delta y_0}{\delta}; \quad a_2 = \frac{\Delta^2 y_0}{2!\delta^2}; \ \ldots; \ a_n = \frac{\Delta^n y_0}{n!\delta^n}.$$

Sustituyendo estos valores en la fórmula (6-66), obtenemos la fórmula de interpolación de Newton:

$$q(x) \approx p(x) = y_0 + \Delta y_0 \left(\frac{x - x_0}{\delta}\right) + \frac{\Delta^2 y_0}{2!}\left(\frac{x - x_0}{\delta}\right)\left(\frac{x - x_1}{\delta}\right) + \ldots$$

$$\ldots + \frac{\Delta^n y_0}{n!}\left(\frac{x - x_0}{\delta}\right)\left(\frac{x - x_1}{\delta}\right) \ldots \left(\frac{x - x_{n-1}}{\delta}\right). \quad (6\text{-}67)$$

Esta fórmula se escribe corrientemente de otra manera, haciendo

$$\frac{x - x_0}{\delta} = u \quad \text{o} \quad x = x_0 + \delta u. \quad (6\text{-}68)$$

Tomando en consideración (6-68), la fórmula de Newton adquiere la forma de

$$p(x) \approx p(x_0 + \delta u) = y_0 + u\,\Delta y_0 +$$

$$+ \frac{u(u-1)}{2!}\Delta^2 y_0 + \ \ldots \ + \frac{u(u-1)(u-2)\ldots(u-n+1)}{n!}\Delta^n y_0. \quad (6\text{-}69)$$

En diversos casos para interpolar es cómodo emplear la fórmula

$$p(x) = \sum_{j=0}^{n} y_j \prod_{\substack{l=0 \\ l \neq j}}^{n} \frac{(x - x_l)}{(x_j - x_l)}. \quad (6\text{-}70)$$

Es evidente la validez de esta fórmula porque, en primera, ella representa un polinomio de x, y, en segunda, para $x = x_k$, $k = 0, 1, \ldots, n$ los valores $p(x_k)$ resultan iguales a y_k.

Capítulo séptimo

PROGRAMACIÓN LINEAL

7-1. ENUNCIADO DEL PROBLEMA DE LA PROGRAMACION LINEAL

a) Definiciones fundamentales

Como se indicó en el capítulo 6, el término "programación lineal" está relacionado con la investigación y resolución del problema siguiente.

Está dado un sistema de m ecuaciones lineales independientes con n incógnitas $x_1, \ldots, x_n$ llamado *sistema de limitaciones* del problema de programación lineal:

$$\left.\begin{array}{l} a_{11}x_1 + \ldots + a_{1n}x_n = b_1; \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ a_{m1}x_1 + \ldots + a_{mn}x_n = b_m. \end{array}\right\} \tag{7-1}$$

El rasgo característico del problema dado es que el número de ecuaciones es menor que el de incógnitas, o sea, $m < n$. Se requiere hallar los valores no negativos de las variables ($x_i \geqslant 0$, $i = 1, \ldots, n$) que satisfacen las ecuaciones (7-1) y minimizan la función de objetivo:

$$q = c_0 + c_1x_1 + \ldots + c_nx_n, \tag{7-2}$$

frecuentemente denominada *forma lineal.*

Hágamos algunas aclaraciones a propósito de este problema. El sistema de ecuaciones (7-1) para el caso en que el número de ecuaciones es igual al número de incógnitas ($m = n$) se estudia en el álgebra ordinaria. Si con esto la determinante del sistema no es igual a cero, el sistema tiene una sola solución y para hallarla hay procedimientos bien elaborados. Si el número de ecuaciones es menor que el de incógnitas ($m < n$), el sistema de ecuaciones (7-1) tiene un conjunto infinito de soluciones, es decir, hay un conjunto infinito de colecciones de variables x_i ($i = 1, \ldots, n$) que satisfacen las ecuaciones (7-1). Vamos a llamar solución a cada colección de variables x_i ($i = 1, \ldots, n$) que satisface el sistema de ecuaciones (7-1).

No obstante, se ha impuesto a las variables x_i una limitación adicional que consiste en que estas variables deben ser no negativas ($x_i \geqslant 0$). En el caso general hay un conjunto infinito de soluciones que satisfacen esta condición adicional. Vamos a llamar solución admisible a cualquier solución del sistema (7-1) con valores no negativos de las variables. De este modo, la esencia del

problema de la programación lineal consiste en elegir del conjunto de soluciones admisibles una solución y precisamente tal que minimiza la forma lineal (7-2).

A veces en un problema de programación lineal todas o algunas de las ecuaciones del tipo (7-1) tienen la forma de desigualdades. Así, en lugar de la ecuación

$$a_{j1}x_1 + \ldots + a_{jn}x_n = b_j \tag{7-3}$$

puede formar parte del sistema (7-1) una desigualdad del tipo

$$a_{j1}x_1 + \ldots + a_{jn}x_n \leqslant b_j, \tag{7-4}$$

o bien

$$a_{j1}x_1 + \ldots + a_{jn}x_n \geqslant b_j. \tag{7-5}$$

Pero es fácil transformar estas desigualdades en ecuaciones, introduciendo la variable adicional $x_{n+j} \geqslant 0$ de modo que, según el signo de la desigualdad, tenga lugar una de las dos expresiones

$$\left.\begin{aligned} a_{j1}x_n + \ldots + a_{jn}x_n + x_{n+j} = b_j; \\ a_{j1}x_n + \ldots + a_{jn}x_n - x_{n+j} = b_j. \end{aligned}\right\} \tag{7-6}$$

Este cambio provoca simplemente el aumento del número de variables lo que no altera la esencia del problema.

A veces no existe el problema de minimizar la forma lineal (7-2) sino de maximizarla. Este problema se reduce al anterior cambiando el signo de la expresión para q. Por cuanto en lo sucesivo van a encontrarse problemas de ambos tipos, acordemos designar el valor de la función de objetivo por q, si hay que minimizarla y por q', si debe ser maximizada.

Antes de buscar la solución de la ecuación (7-1) que satisface todas las exigencias impuestas ($x_i \geqslant 0$, $i = 1, \ldots, n$; $q = \text{mín}$), tratemos de hallar alguna solución de la ecuación (7-1). Por cuanto el número de variables n en este sistema es mayor que el de ecuaciones m, una de las soluciones posibles puede hallarse si hacemos igual a cero $n - m$ variables cualesquiera. El sistema de m ecuaciones con m incógnitas puede resolverse por los métodos habituales del álgebra. Es verdad que para que el sistema de m ecuaciones con m incógnitas tenga solución, es necesario que no se reduzca a cero la determinante compuesta con los coeficientes de las incógnitas. Si esta condición no se cumple, pueden igualarse a cero otras $n - m$ variables. La solución obtenida con esto se denomina solución *de base*. Ahora podemos introducir la terminología que se emplea ampliamente en los problemas de programación lineal.

Se denomina *base* calquier colección de m variables tales que la determinante compuesta con los coeficientes de estas variables no es igual a cero. Estas m variables se denominan variables *de base* (con respecto a la base dada). Las $n - m$ variables restantes se denominan variables *no de base* o *libres*. En cada sistema de

ecuaciones concreto (7-1) pueden existir varias bases diferentes con distintas variables de base.

Si hacemos iguales a cero todas las variables libres y resolvemos el sistema obtenido de m ecuaciones con m incógnitas, logramos llegar a la solución *de base.* Pero entre las diferentes soluciones de base van a haber algunas que dan valores negativos de ciertas variables. Estas soluciones de base contradicen la hipótesis del problema y son *inadmisibles.*

Constituye una solución de base *admisible* la que es admisible simultáneamente, o sea, la solución de base que da valores no negativos de las variables de base. Las soluciones de base admisibles son las más sencillas de las soluciones admisibles del sistema (7-1). No obstante, en la solución del problema se implanta una condición adicional: la forma lineal (7-2) debe adquirir el valor mínimo con la solución hallada. Con esta condición adicional se complica la solución del problema, no obstante, el concepto de solución de base admisible juega un papel importantísimo al hallar la solución completa del problema.

b) Ejemplos de problemas de programación lineal

La programación lineal surgió debido al estudio de los asuntos relacionados con la búsqueda de las variantes más ventajosas durante la resolución de diversos problemas de planificación y producción. En estos problemas existen una gran libertad de variación de diferentes parámetros y una serie de condiciones limitadoras. Se requiere hallar tales valores de los parámetros que desde algún punto de vista sean los mejores. Se catalogan entre dichos problemas los de la búsqueda del método más racional de utilización de las materias primas y los materiales, la determinación de los regímenes de producción más ventajosos, el aumento de la eficacia de funcionamiento de los medios de transporte, etc. Durante el mando automático del proceso de producción estas tareas deben solucionarse automática y continuamente. Por eso, el conocimiento de los métodos de resolución de problemas similares que es necesario para cualquier ingeniero, se convierte sumamente valioso para el especialista en automática.

Aclararemos con una serie de ejemplos la esencia de los problemas de programación lineal.

Ejemplo 7-1. Problema de la utilización de recursos. Para realizar l procesos tecnológicos diferentes $T_1, \ldots, T_l$ una fábrica necesita m tipos de recursos $S_1, \ldots, S_m$ (materias primas, combustible, materiales, herramientas, etc). Las reservas de recursos de cada tipo son limitadas e iguales a $b_1, \ldots, b_m$. Se conoce el gasto de recursos por unidad de producción en cada proceso tecnológico. Se requiere determinar qué cantidad de producción de cada tipo hay que elaborar para que sea máximo el ingreso debido a la realización de esta producción.

Designemos por a_{ij} el gasto de los recursos del tipo S_i por unidad de producción del tipo T_j y por c_j, el ingreso debido a la realización de la unidad

de producción del tipo T_j. Representemos todos los datos existentes en forma de la tabla 7-1 suponiendo para ser concretos que $l = 3$ y $m = 4$

Tabla 7-1

Anotación de los datos iniciales en el problema de la utilización de recursos

Tipos de recursos	Gasto de recursos por unidad de producción			Reservas de recursos
	T_1	T_2	T_3	
S_1	a_{11}	a_{12}	a_{13}	b_1
S_2	a_{21}	a_{22}	a_{23}	b_2
S_3	a_{31}	a_{32}	a_{33}	b_3
S_4	a_{41}	a_{42}	a_{43}	b_4
Ingresos por realización de la unidad de producción	c_1	c_2	c_3	—

Designemos por x_j el número de unidades de producción elaborada del tipo T_j. Son limitaciones en este problema los requerimientos de que el gasto de los recursos del tipo S_i ($i = 1, \ldots, m$) para la elaboración de todos los tipos de producción no sobrepase de las reservas existentes:

$$\sum_{j=1}^{l} a_{ij}x_j \leqslant b_j, \quad i = 1, \ldots, m. \tag{7-7}$$

Estas limitaciones se convierten fácilmente en ecuaciones, introduciendo las variables $x_{l+i} \geqslant 0$, que significan los recursos no utilizados del tipo S_i. Con esto, en lugar de (7-7) obtenemos:

$$\sum_{j=1}^{l} a_{ij}x_j + x_{l+i} = b_i, \quad i = 1, \ldots, m. \tag{7-8}$$

La magnitud del ingreso debido a la realización de la producción elaborada será igual a

$$q' = \sum_{j=1}^{l} c_j x_j. \tag{7-9}$$

El plan óptimo de elaboración de la producción será la solución no negativa del sistema de ecuaciones (7-8) con la que la función de objetivo (7-9) sea máxima.

Ejemplo 7-2. Problema de la distribución por empresas de la producción elaborada. El plan de la rama industrial prevé durante el tiempo T la elaboración de los tipos de producción siguientes:

A_1 en cantidad de N_1 unidades;
A_2 en cantidad de N_2 unidades;
. . . .
A_l en cantidad de N_l unidades.

Estos tipos de producción pueden elaborarse en r empresas similares $E_1, \ldots, E_r$. Supongamos que ninguna de las empresas puede elaborar simultáneamente varios tipos de productos. Además, están dados:

a_{ij}, la cantidad del producto A_i elaborada en la empresa E_j en la unidad de tiempo;

b_{ij}, el costo de la unidad del producto del tipo A_i elaborada en la empresa E_j;

x_{ij}, el tiempo de funcionamiento de la empresa E_j para elaborar el producto A_i.

Se requiere hallar los valores x_{ij} con los que será mínimo el costo de la producción elaborada.

Limitaciones:

1) el tiempo de funcionamiento de cada empresa no debe exceder de T

$$\sum_{i=1}^{l} x_{ij} \leqslant T \quad j = 1, \ldots, r; \tag{7-10}$$

2) la cantidad del producto elaborado debe corresponder a la nomenclatura

$$\sum_{j=1}^{r} a_{ij} x_{ij} = N_i, \quad i = 1, \ldots, l. \tag{7-11}$$

La función de objetivo va a representar el costo total de la producción elaborada. Si se toma en cuenta que la magnitud $a_{ij}b_{ij}x_{ij}$ representa el costo de la parte de la producción A_i elaborada por la empresa E_j, el costo total de la producción elaborada es

$$q = \sum_{j=1}^{r} \sum_{i=1}^{l} a_{ij} b_{ij} x_{ij}. \tag{7-12}$$

Según las condiciones del problema esta magnitud debe minimizarse al cumplirse las limitaciones (7-10) y (7-11).

Ejemplo 7-3. Problema del transporte. En los puntos $P_1, \ldots, P_l$ hay un peso homogéneo en cantidades $a_1, \ldots, a_l$. Es necesario transportarlo a los puntos $Q_1, \ldots, Q_r$ en cantidades $b_1, \ldots, b_r$ de suerte que el costo total del transporte sea mínimo. Con esto se presupone que la cantidad de carga requerida es igual a las reservas existentes

$$\sum_{i=1}^{l} a_i = \sum_{j=1}^{r} b_j. \tag{7-13}$$

Designemos por x_{ij} lan cantidad de carga transportada del punto P_i al punto Q_j, y por c_{ij}, el costo de transporte de la unidad de esta carga. En el problema hay las limitaciones siguientes:

1) la cantidad de carga enviada del punto P_i a todos los puntos de destinación debe ser igual a las reservas existentes a_i:

$$\sum_{j=1}^{r} x_{ij} = a_i, \quad i = 1, \ldots, l; \tag{7-14}$$

2) la cantidad de carga que llega a Q_j desde todos los puntos de envío debe ser igual a la necesidad b_j:

$$\sum_{i=1}^{l} x_{ij} = b_j, \quad j = 1, \ldots, r. \tag{7-15}$$

La función de objetivo determina el costo total de transporte de todas las cargas

$$q = \sum_{j=1}^{r} \sum_{i=1}^{l} c_{ij} x_{ij}. \tag{7-16}$$

Ejemplo 7-4. Problema de selección de la versión óptima de un equipo. Se necesita proyectar un dispositivo calculador numérico que debe efectuar sucesivamente r operaciones matemáticas y conforme a esto consta de r bloques en serie. Hay l versiones diferentes de ejecución de cada bloque: con tubos electrónicos, elementos semiconductores, elementos de ferrita y transistores, micromódulos, etc. Se han establecido limitaciones con respecto al costo máximo (X), las dimensiones exteriores máximas (Y) y el tiempo máximo de realización de las operaciones (Z). Se requiere elegir la versión más ventajosa desde el punto de vista de las exigencias planteadas.

Designemos por x_i, y_i, z_i, respectivamente, el costo, las dimensiones y el tiempo de realización de una operación para el bloque i. Entonces, las limitaciones existentes pueden escribirse en la forma

$$\sum_{i=1}^{r} x_i \leqslant X, \quad \sum_{i=1}^{r} y_i \leqslant Y, \quad \sum_{i=1}^{r} z_i = Z. \tag{7-17}$$

Las magnitudes x_i, y_i y z_i dependen de la versión de ejecución del bloque, es decir, son elementos de los conjuntos siguientes:

$$x_i \in \{x_{i1}, \ldots, x_{il}\}; \quad y_i \in \{y_{i1}, \ldots, y_{il}\};$$
$$z_i \in \{z_{i1}, \ldots, z_{il}\}, \tag{7-18}$$

donde las magnitudes x_{ij}, y_{ij}, z_{ij} significan, respectivamente, el costo, las dimensiones exteriores y el tiempo de realización de una operación para el bloque i con la versión de ejecución j.

Si c_1, c_2, c_3 son los coeficientes que caracterizan el valor relativo de la disminución del costo, las dimensiones exteriores y el tiempo de realización de las operaciones, entonces, la condición de la versión óptima del equipo se escribirá en la forma

$$q = c_1 \sum_{i=1}^{r} x_i + c_2 \sum_{i=1}^{r} y_i + c_3 \sum_{i=1}^{r} z_i = \min. \tag{7-19}$$

La particularidad del problema dado que lo diferencia del de programación lineal enunciado inicialmente es que aunque las limitaciones (7-17) y la función de objetivo (7-19) son lineales, no obstante, las variables x_i, y_i, z_i no pueden tomar cualesquiera valores no negativos, sino únicamente los de los conjuntos finitos (7-18). Por eso, los métodos corrientes de programación lineal no son aplicables al problema dado y éste debe resolverse por los métodos de programación discreta (de números enteros).

c) Interpretación geométrica del problema de programación lineal

Para tener una idea más completa del problema de la programación lineal daremos la interpretación geométrica del problema siguiente. Se tiene el sistema de ecuaciones

$$\left.\begin{aligned} -2x_1 + x_2 + x_3 &= 2; \\ x_1 - 2x_2 + x_4 &= 2; \\ x_1 + x_2 + x_5 &= 5 \end{aligned}\right\} \tag{7-20}$$

y la función de objetivo

$$q = x_2 - x_1. \tag{7-21}$$

Se pretende hallar los valores no negativos de las variables que satisfacen las ecuaciones (7-20) y que minimizan la función de objetivo (7-21). Nos será más cómodo dar las reglas de solución del problema de programación lineal para el caso de maximización de la función de objetivo, lo que es fácil hacer tomando en calidad de función de objetivo la expresión

$$q' = -q = x_1 - x_2. \tag{7-22}$$

En el ejemplo considerado el número de ecuaciones es $m = 3$ y el de incógnitas, $n = 5$, así que hay $m = 3$ variables de base y $n - m = 2$ variables libres. El hecho de que sólo hay dos variables libres hace posible ilustrar geométricamente la solución del problema en el espacio binario, es decir, en el plano.

El sistema de tres ecuaciones (7-20) puede resolverse con respecto a tres variables, por ejemplo, x_3, x_4 y x_5 expresándolas por medio de x_1 y x_2; al hacerlo, obtenemos:

$$\left.\begin{aligned} x_3 &= 2 + 2x_1 - x_2; \\ x_4 &= 2 - x_1 + 2x_2; \\ x_5 &= 5 - x_1 - x_2. \end{aligned}\right\} \tag{7-23}$$

Por condición del problema de programación lineal las variables sólo pueden tomar valores no negativos, es decir, el dominio de valores admisibles de las variables será el dominio determinado por las condiciones

$$x_i \geqslant 0, \quad i = 1,\ 2,\ 3,\ 4,\ 5. \tag{7-24}$$

Cada una de las desigualdades (7-24) define cierto semiplano en el plano (x_1, x_2). Así pues, la desigualdad $x_1 \geqslant 0$ define el semiplano superior y la desigualdad $x_3 \geqslant 0$, el semiplano que se encuentra por un lado de la recta $2 + 2x_1 - x_2 = 0$ y precisamente aquel que contiene el origen de coordenadas, lo que se comprueba con facilidad sustituyendo en la primera de las desigualdades (7-23) las coordenadas del punto (0, 0). El dominio correspondiente a $x_3 < 0$ se llama *prohibido* y es útil marcarlo mediante rayado como se muestra en la figura 7-1.

En la figura 7-2 se muestran trazados similares para todas las x_i. En esta figura las rectas correspondientes a la condición $x_i = 0$ ($i = 1, 2, 3, 4, 5$) están marcadas las cifras. Los dominios prohibidos correspondientes a la condición $x_i < 0$ están marcados por rayado. Como se ve en la figura 7-2 el dominio de valores admisibles de las variables x_1 y x_2 es la zona obscura que representa el polígono *oabcd*. Tiene importancia apuntar que el polígono de soluciones admisibles es convexo ya que constituye la intersección de los dominios convexos definidos por las condiciones $x_i \geqslant 0$.

El trazado realizado permite dar la interpretación geométrica de la solución de base. Como que cada recta en la figura 7-2 corresponde a la anulación de una de las variables, en los puntos de intersección de dos rectas van a convertirse en cero dos, es decir, $n - m$ variables. Pero $n - m$ es el número de variables libres cuya anulación corresponde a la solución de base. De esta manera, los puntos de intersección de las rectas $x_i = 0$ $(i = 1, \ldots, 5)$ determinan las soluciones de base de los problemas de programación lineal.

Entre las soluciones de base hay algunas que no pertenecen al dominio de soluciones admisibles. Estas son soluciones de base inadmisibles. Sólo pertenecen al dominio de soluciones admisibles

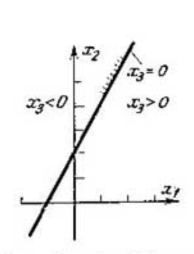


Fig. 7-1. Semiplano $2 + 2x_1 - x_2 \geqslant 0$

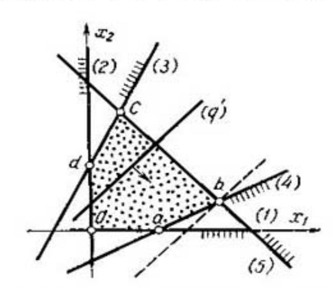


Fig. 7-2. Polígono de soluciones admisibles

aquellos puntos de intersección de las rectas $x_i = 0$ que son vértices del polígono de soluciones admisibles. Por consiguiente, los vértices del polígono de soluciones admisibles corresponden a soluciones de base admisibles.

Examinemos ahora la condición de maximización de la función de objetivo (7-22). Geométricamente, con cualquier q' la expresión (7-22) define una recta trazada formando ángulo de 45° con el eje de las abscisas, con esto, al incremento de q' corresponde el desplazamiento de la recta en la dirección de la flecha mostrada en la figura 7-2. Esta recta estará compatible con la solución admisible del problema sólo en el caso cuando tenga puntos comunes con el dominio de soluciones admisibles. El valor máximo de q' se obtiene con recta en la posición extrema cuando ella se convierte en recta de apoyo hacia el dominio de soluciones admisibles (línea de trazos en la figura 7-2). No obstante, la recta de apoyo que va al polígono convexo pasa, sin falta, a través de uno siquiera de sus vértices que, como vimos, corresponden a las soluciones de base admisibles.

De esta manera, llegamos a la deducción importantísima de que la solución de un problema de programación lineal que convierte

en máximo la función de objetivo q', se halla sin aflta entre las soluciones *de base admisibles.*

Esta deducción se generaliza fácilmente para el caso de m y n arbitrarios, cuando $n - m > 2$. En este caso las condiciones $x_i \geqslant 0$ $(i = 1, \ldots, n)$ determinan los semiespacios en el espacio multidimensional cuya intersección definirá el dominio de soluciones admisibles en forma de poliedro convexo. Los vértices de este poliedro corresponderán a las soluciones de base admisibles. La expresión para la función de objetivo define el hiperplano en el espacio multidimensional considerado, cuyo hiperplano con la condición $q' =$ máx será de apoyo al poliedro de soluciones admisibles, es decir, pasarán sin falta por uno siquiera de los vértices. En este caso la solución del problema de programación lineal va a hallarse también entre las soluciones *de base admisibles.*

La deducción obtenida permite determinar el procedimiento de resolución del problema de programación lineal. Por cuanto la sodición será de apoyo al poliedro de soluciones admisibles, cuyo número es finito, pueden hallarse todas las soluciones de base admisibles y calcularse el valor de q' para cada una de ellas. La solución definitiva es aquella de las soluciones halladas para la cual el valor q' será máximo.

Este procedimiento de resolución del problema aunque factible es sumamente engorroso ya que el número de soluciones de base admisibles puede ser muy grande. No obstante, existen métodos racionales de selección consecutiva de las soluciones de base, que permiten examinar no todas las soluciones de base admisibles sino un número mínimo de ellas. Uno de los métodos más difundidos de esta selección es el llamado método símplex a cuyo estudio estará dedicado el párrafo siguiente.

7-2. RESOLUCION DEL PROBLEMA DE PROGRAMACION LINEAL

a) Algebra del método símplex

La esencia del método símplex consiste en lo siguiente. Ante todo se halla alguna solución de base admisible. Esta puede encontrarse tomando $n - m$ variables cualesquiera como libres, igualándolas a cero y resolviendo el sistema de ecuaciones obtenido (7-1). Si con esto resultan negativas algunas de las variables de base, es necesario elegir otras variables libres, o sea, pasar a una nueva base.

Después de haberse encontrado la solución de base admisible se comprueba si no se ha alcanzado ya el máximo de la función de objetivo q'. Si no es así, se busca una nueva solución de base, pero no cualquiera sino aquella que aumenta el valor de la función de objetivo q'. Después se repite el proceso. Por cuanto en calidad

de nueva solución de base admisible se elige sólo aquella que aumenta el valor de la función de objetivo, el método dado brinda la posibilidad de examinar un mínimo de soluciones de base admisibles y conduce con bastante rapidez al objetivo. Analicemos este método más detalladamente con el ejemplo examinado anteriormente.

Recurramos al sistema de ecuaciones (7-20) para el cual se requiere hallar una solución negativa que maximice la función de objetivo (7-22). Por cuanto $n - m = 2$, pueden tomarse en calidad de variables libres dos cualesquiera, por ejemplo x_1 y x_2. Igualándolas a cero, de (7-20) hallamos la solución de base $x_1 = x_2 = 0$, $x_3 = 2$, $x_4 = 2$, $x_5 = 5$ que es admisible y con la cual $q' = x_1 - x_2 = 0$.

La comprobación de si con la solución hallada ha alcanzado o no el máximo de la función de objetivo puede efectuarse buscando una nueva solución de base con la cual dicha función será mayor. Para el paso a la nueva solución de base admisible hay que hacer variable de base una de las variables libres x_1 o x_2. Con esto ella va a ser diferente de cero, es decir, crecerá. Por consiguiente, si cualquiera de las variables libres forma parte de la expresión para la función de objetivo con signo «+», vale decir, al aumentar ésta crece la función de objetivo, entonces no se ha alcanzado el máximo de la función de objetivo y hay que convertir la variable dada en variable de base haciéndola diferente de cero.

Sin embargo, en caso de crecer la variable libre comienzan a disminuir algunas de las variables de base. Como que los valores negativos de las variables son inadmisibles, hay que tomar en calidad de nueva variable libre aquella de las variables de base que se convierte en cero antes que las demás.

En el ejemplo examinado forma parte de la expresión para la función de objetivo (7-22) con signo «+» la variable libre x_1. Esto significa que la función de objetivo no ha alcanzado el máximo y debe hacerse variable de base la x_1. Para determinar la nueva variable libre expresaremos las variables de base x_3, x_4, x_5 por medio de las variables libres reduciendo las ecuaciones (7-20) a la forma (7-23). En estas ecuaciones vemos que con $x_2 = 0$ el aumento de x_1 no provoca la disminución de x_3 pero hace decrecer a x_4 y x_5 de modo que siendo $x_1 = 2$ obtenemos $x_4 = 0$ y $x_5 = 3 > 0$. De este modo la variable x_4 debe hacerse libre lo que nos lleva a una nueva base x_1, x_3, x_5.

Para continuar la resolución, solucionaremos el sistema (7-20) respecto a las nuevas variables de base, reduciéndolo a la forma

$$\left.\begin{aligned} x_1 &= 2 + 2x_2 - x_4; \\ x_3 &= 6 + 3x_2 - 2x_4; \\ x_5 &= 3 - 3x_2 + x_4. \end{aligned}\right\} \qquad (7\text{-}25)$$

Asimismo expresaremos la función de objetivo por medio de las nuevas variables libres x_2 y x_4:

$$q' = 2 + x_2 - x_4. \tag{7-26}$$

Repitiendo los razonamientos anteriores llegamos a deducir que no se ha alcanzado el máximo de la función de objetivo y hay que pasar la variable libre x_2 a la de base y la variable de base x_5, a la libre. Con esto, las variables x_1, x_2, x_3 forman una nueva base.

Resolviendo las ecuaciones (7-20) respecto de las nuevas variables de base, obtenemos:

$$\left.\begin{aligned} x_1 &= 4 - \frac{1}{3}x_4 - \frac{2}{3}x_5; \\ x_2 &= 1 + \frac{1}{3}x_4 - \frac{1}{3}x_5; \\ x_3 &= 9 - x_4 - x_5. \end{aligned}\right\} \tag{7-27}$$

La función de objetivo toma entonces la forma:

$$q' = 3 - \frac{2}{3}x_4 - \frac{1}{3}x_5. \tag{7-28}$$

En la expresión (7-28) vemos que, al aumentar las variables libres x_4 y x_5, disminuye el valor de q'. Por tanto, con la base dada se alcanza el máximo de la función de objetivo q' y la solución del problema considerado será la colección de variables siguientes:

$$x_4 = x_5 = 0, \quad x_1 = 4, \quad x_2 = 1, \quad x_3 = 9.$$

Con esto $q' = -q = 3$, $q = -3$.

El método de resolución del problema de programación lineal estudiado posee la deficiencia de que acarrea voluminosas transformaciones de una forma a otra del sistema de ecuaciones lineales. Puede lograrse una considerable simplificación de las transformaciones si presentamos las ecuaciones en forma de tablas que contengan los coeficientes de las variables. Con esto, el paso de un sistema de ecuaciones al otro se reduce a calcular nuevamente los coeficientes en las tablas lo que se realiza según reglas puramente formales que, además, están bien adaptadas para resolverlas con calculadoras electrónicas.

b) Método tabular de búsqueda de la solución óptima

Al emplear el método tabular es conveniente introducir una forma especial de notación para la ecuación (7-1) y la función de objetivo (7-2). Designemos por x'_i, $i = 1, \ldots, m$ las variables de base y por x''_j, $j = 1, \ldots, (n - m)$, las variables libres. Habiendo expresado la función de objetivo y las variables de base por medio de variables libres, enunciaremos el problema de programación lineal de la suerte siguiente:

maximizar

$$q' = -q = \alpha_{00} - \sum_{j=1}^{n-m} \alpha_{0j} x''_j \quad (7\text{-}29)$$

a condiciones de que

$$x'_i = \alpha_{i0} - \sum_{j=1}^{n-m} \alpha_{ij} x''_j; \quad i = 1, \ldots, m;$$

$$x'_i \geqslant 0; \quad x''_j \geqslant 0. \quad (7\text{-}30)$$

Con tal forma de notación el problema puede representarse mediante la siguiente matriz de coeficientes de las variables:

$$\begin{array}{c} \\ q' \\ x'_1 \\ \cdots \\ x'_m \end{array} \begin{array}{c} 1 \quad -x''_1 \quad \ldots \quad -x''_{n-m} \\ \left\| \begin{array}{cccc} \alpha_{00} & \alpha_{01} & \ldots & \alpha_{0\,(n-m)} \\ \alpha_{10} & \alpha_{11} & \ldots & \alpha_{1\,(n-m)} \\ \cdot & \cdot & \cdot & \cdot \\ \alpha_{m0} & \alpha_{m1} & \ldots & \alpha_{m\,(n-m)} \end{array} \right\| \end{array} . \quad (7\text{-}31)$$

Por el aspecto de los coeficientes de la matriz (7-31) es fácil juzgar si es admisible la solución de base hallada y, en caso de ser admisible, si será o no óptima. En efecto, advirtiendo que la columna de coeficientes α_{i0}, $i \neq 0$ representa la solución de base correspondiente a la base $x'_1, \ldots, x'_m$, y la línea de coeficientes α_{0j}, $j \neq 0$, los coeficientes de variables libres en la expresión para q' tomados con el signo opuesto, llegamos a la conclusión de que la solución de base correspondiente a la base $x'_1, \ldots, x'_m$ es admisible si $\alpha_{i0} \geqslant 0$, $i \neq 0$. Si, además, $\alpha_{0j} \geqslant 0$, $j \neq 0$, esta solución de base es asimismo la óptima. También, es evidente que con la solución de base óptima el coeficiente α_{00} brinda el valor $q' = q'_{máx} = -q_{mín}$.

Comencemos la solución del problema buscando cualquier solución de base admisible que no sea óptima en caso general. Representemos esta solución de base mediante la tabla de coeficientes compilada a modo de la matriz (7-31). Para el paso a la mejor solución de base es necesario transformar la matriz de los coeficientes. El método de esta transformación se formulará a manera de sistema de reglas que ilustraremos con el ejemplo del problema resuelto en el párrafo anterior. No vamos a dar una demostración rigurosa a estas reglas. Sólo hacemos hincapié en que ellas corresponden exactamente a aquellas transformaciones del sistema de ecuaciones que se realizaron en el ejemplo precedente y pueden comprobarse fácilmente comparando las transformaciones correspondientes.

Supongamos que el problema de programación lineal está planteado en forma del sistema de ecuaciones (7-20) y de la función de objetivo (7-22) que hace falta maximizar. Tomando x_1 y x_2 como

variables libres, reducimos este sistema de ecuaciones y la función de objetivo a las formas (7-29) y (7-30):

$$\left.\begin{aligned} q' &= 0 - (-x_1 + x_2); \\ x_3 &= 2 - (-2x_1 + x_2); \\ x_4 &= 2 - (x_1 - 2x_2); \\ x_5 &= 5 - (x_1 + x_2). \end{aligned}\right\} \qquad (7\text{-}32)$$

Representemos la matriz de los coeficientes en forma de la tabla 7-2, *a* con cuadros de dimensiones suficientemente grandes, en cuyo ángulo superior izquierdo anotaremos los coeficientes α_{ij} de las ecuaciones (7-32).

Tabla 7-2

Transformaciones consecutivas de las tablas de coeficientes al resolver un problema de programación lineal

a) (en cada casilla: ángulo superior izquierdo / ángulo inferior derecho)

	1	$-x_1$	$-x_2$
q'	0 / 2	−1 / 1	1 / −2
x_3	2 / 4	−2 / 2	1 / −4
x_4	2 / 2	[1] / 1	−2 / −2
x_5	5 / −2	5 / −1	1 / 2

b)

	1	$-x_4$	$-x_2$
q'	2 / 1	1 / $-\frac{1}{3}$	−1 / $\frac{1}{3}$
x_3	6 / 3	2 / −1	−3 / 1
x_1	2 / 2	1 / $\frac{2}{3}$	−2 / $\frac{2}{3}$
x_5	3 / 1	−1 / $-\frac{1}{3}$	[3] / $\frac{1}{3}$

c)

	1	$-x_4$	$-x_5$
q'	3	$\frac{2}{3}$	$\frac{1}{3}$
x_3	9	1	1
x_1	4	$\frac{1}{3}$	$\frac{2}{3}$
x_2	1	$-\frac{1}{3}$	$\frac{1}{3}$

Comprobemos si no se ha encontrado ya la solución óptima cuya condición es $\alpha_{0j} \geqslant 0$, $j \neq 0$. Por cuanto α_{01} (coeficiente de $-x_1$ en la expresión para q') es negativo, la solución óptima no se ha encontrado y la variable x_1 debe hacerse variable de base. Destaquemos la columna correspondiente a la variable x_1 por medio de rayas dobles. Si los coeficientes α_{0j} de algunas variables libres resultan negativos entonces puede convertirse en variable de base cualquiera de ellas.

Determinemos ahora cuál de las variables de base debe hacerse libre. Evidentemente aquella que con más prontitud se vuelve cero

al aumentar x_1. Esta será la variable de base x_i para la cual el coeficiente en la columna marcada $\alpha_{i1} > 0$ y la relación α_{i0}/α_{i1} es la menor. Dicha variable es la de base x_4 con $\alpha_{40} = 2$ y $\alpha_{41} = 1$. También marcaremos con líneas dobles la fila correspondiente a la variable de base x_4.

Llamaremos *coeficiente general* al coeficiente $\lambda = \alpha_{41}$ que se encuentra en el ángulo superior izquierdo del cuadro en la intersección de la fila y la columna marcadas. En nuestro caso $\lambda = 1$ (encerrado en un marco).

Ahora deben llenarse los ángulos inferiores de cada cuadro. Lo hacemos según las reglas siguientes:

1) en el cuadro, en la intersección de la fila y la columna marcadas escribimos $1/\lambda$;

2) en los cuadros de la fila marcada escribimos los coeficientes superiores, multiplicados por λ (los coeficientes superiores a excepción del coeficiente general se distinguen con tipos gruesos);

3) en los cuadros de la columna marcada escribimos los coeficientes superiores multiplicados por $-\lambda$ (los coeficientes inferiores a excepción del cuadro con el coeficiente general se distinguen con tipos gruesos);

4) en los cuadros restantes escribimos el producto de los coeficientes distinguidos con tipos gruesos en cuya intersección se encuentra el cuadro dado.

Después pasamos a llenar la tabla 7-2, *b*, que difiere de la 7-2, *a* en que la variable libre marcada con x_1 se convierte en variable de base y la variable de base marcada con x_4 se volvió libre. Los ángulos superiores izquierdos de los cuadros de la tabla 7-2, *b* se llenan según las reglas siguientes:

1) la fila y la columna correspondientes a las nuevas variables libre y de base se llenan con los coeficientes inferiores de la fila y la columna marcadas de la tabla 7-2, *a*;

2) en los cuadros restantes se anotan las sumas de los coeficientes quese hallan en los cuadros correspondientes de la tabla 7-2, *a*.

La tabla 7-2, *b* llenada de esta forma corresponde a la matriz de los coeficientes (7-31) con la nueva base x_1, x_3, x_5. En lo sucesivo se repite todo el proceso.

Por cuanto en la tabla 7-2, *b* el coeficiente $\alpha_{02} < 0$, no se ha encontrado la solución óptima. Luego, según las reglas enunciadas es necesario llenar los ángulos inferiores izquierdos de la tabla 7-2, *b* y pasar a la nueva tabla 7-2, *c* que corresponde a la base x_1, x_2, x_3. En esta tabla los coeficientes α_{0j}, $j \neq 0$ son positivos y ella brinda la solución óptima del problema que hallamos en la columna de los miembros libres:

$$x_1 = 4; \quad x_2 = 1; \quad x_3 = 9; \quad x_4 = x_5 = 0;$$

$$q'_{\text{máx}} = -q_{\text{mín}} = 3; \quad q_{\text{mín}} = -3.$$

c) Problema dual de programación lineal

Volvamos al problema de distribución de recursos examinado en el ejemplo 7-1 y preguntémonos cuál es, desde el punto de vista de la empresa, el valor de los recursos existentes a su disposición. Al solucionar esta cuestión vamos a tener en cuenta que los recursos que la empresa no puede utilizar por completo tienen para ella un valor muy bajo en el sentido de que la misma no va a estar de acuerdo en sufragar incluso gastos pequeños para aumentar las reservas de dichos recursos. Así pues, el equipo costoso que no se utiliza en el proceso tecnológico posee valor nulo para la empresa. Evidentemente tendrán el mayor valor los recursos que limitan en mayor grado la producción y, por ende, los ingresos de la empresa en el aumento de las reservas de cuyos recursos la empresa está dispuesta a sufragar gastos considerables.

Por eso puede considerarse que cada tipo de recurso posee cierto "precio oculto" que determina el valor del recurso dado para la empresa desde el punto de vista de los ingresos por la realización de la producción y que depende de la reserva existente de este recurso y de las necesidades de él para realizar la producción.

Si la empresa, por motivos cualesquiera, se limita a un solo proceso tecnológico que requiere grandes gastos de cierto recurso cuyas reservas están limitadas, el precio oculto de este recurso será alto. No obstante, los precios ocultos establecidos en concordancia con este proceso tecnológico no serán los mejores ya que la introducción de otros procesos tecnológicos permitirá utilizar más racionalmente todas las reservas de recursos. Luego, existen precios ocultos óptimos, que corresponden a los ingresos máximos de la empresa, es decir, a la distribución óptima de los recursos. Como vemos, la determinación de los precios ocultos óptimos resulta estrechamente vinculada al problema de distribución óptima de los recursos, o sea, al problema de programación lineal descrito mediante el sistema de ecuaciones (7-8) y la función de objetivo (7-9). Sin embargo, para determinar los precios ocultos óptimos puede plantearse un problema independiente de programación lineal.

Designemos por u_i el precio oculto de la unidad del recurso S_i. Las magnitudes u_i deben ser tales que el precio oculto de los recursos utilizados en cualquier proceso tecnológico no sea menor que el ingreso obtenido. Empleando las designaciones del ejemplo 7-1 y los datos de la tabla 7-1, escribiremos esta condición en la forma:

$$\sum_{i=1}^{m} a_{ij}u_i \geqslant c_j, \quad j=1, \ldots, l. \qquad (7\text{-}33)$$

Si introducimos las variables $u_{m+j} \geqslant 0$ que representan el exceso del precio oculto de la unidad del producto sobre los ingresos debidos a su realización, el sistema de desigualdades (7-33) se

convierte en el de ecuaciones

$$\sum_{i=1}^{m} a_{ij}u_i - u_{m+j} = c_j, \quad j = 1, \ldots, l. \tag{7-34}$$

Consideraremos precios ocultos óptimos a los que minimizan el costo total de los recursos, es decir, la magnitud

$$q^* = \sum_{i=1}^{m} b_i u_i. \tag{7-35}$$

El sistema de limitaciones (7-33) junto con la función de objetivo (7-35) constituye un nuevo problema de programación lineal que recibió el nombre de problema *dual* respecto al del ejemplo 7-1 que se denomina problema *directo* o *fundamental* de programación lineal.

No es difícil notar que los problemas directo y dual resultan estrechamente vinculados entre sí. Este vínculo se manifiesta en lo siguiente:

si el problema directo es un problema de maximización de la función de objetivo, el problema dual lo es de minimización;

los coeficientes de la función de objetivo en el problema directo son miembros libres en las limitaciones del problema dual;

los miembros libres de las limitaciones del problema directo se vuelven coeficientes de la función de objetivo en el problema dual;

los coeficientes de las variables en las limitaciones del problema dual representan columnas de la tabla de coeficientes del problema directo;

los signos de las desigualdades en las limitaciones se cambian por los opuestos.

Existe una estrecha relación entre las soluciones de los problemas directo y dual de programación lineal. Para establecer esta relación escribiremos las ecuaciones de los problemas directo y dual en otra forma.

En el problema directo tomaremos las variables $x_1, \ldots, x_l$ como libres y lo formularemos del modo siguiente:

maximizar

$$q' = 0 - \sum_{j=1}^{l} (-c_j)\, x_j \tag{7-36}$$

a condiciones de que

$$x_{l+i} = b_i - \sum_{j=1}^{l} a_{ij}x_j, \quad i = 1, \ldots, m. \tag{7-37}$$

A este problema corresponde la matriz del tipo

$$\begin{array}{c} \\ q' \\ x_{l+1} \\ \cdots \\ x_{l+m} \end{array}
\begin{array}{c} 1 \quad -x_1 \ \ldots \ -x_l \\ \left\|\begin{array}{cccc} 0 & -c_1 & \ldots & -c_l \\ b_1 & a_{11} & \ldots & a_{1l} \\ \cdot & \cdot & \cdot & \cdot \\ b_m & a_{m1} & \ldots & a_{ml} \end{array}\right\| \end{array}. \tag{7-38}$$

En el problema dual tomaremos como variables libres $u_1, \ldots$ $\ldots, u_m$ y lo formularemos de la manera siguiente:
minimizar

$$q^* = 0 \sum_{i=1}^{m} b_i u_i \tag{7-39}$$

a condiciones de que

$$u_{m+j} = -c_j + \sum_{i=1}^{m} a_{ij} u_i, \quad j = 1, \ldots, l. \tag{7-40}$$

A este problema corresponde la matriz del tipo

$$\begin{array}{c} \\ q^* \\ u_{m+1} \\ \cdots \\ u_{m+l} \end{array} \begin{array}{c} \begin{array}{cccc} 1 & u_1 & \ldots & u_m \end{array} \\ \left\| \begin{array}{cccc} 0 & b_1 & \ldots & b_m \\ -c_1 & a_{11} & \ldots & a_{m1} \\ \cdot & \cdot & \cdot & \cdot \\ -c_l & a_{1l} & \ldots & a_{ml} \end{array} \right\| \end{array}. \tag{7-41}$$

Como vemos, las columnas de la matriz (7-41) son líneas de la matriz (7-38). Por consiguiente, podemos describir los problemas directo y dual mediante una misma matriz (7-38) en la cual, no obstante, debe estar establecida la siguiente corronspondencia entre las variables directas y duales:

$$\left. \begin{array}{ll} x_j \leftarrow \rightarrow -u_{m+j}, & j = 1, \ldots, l; \\ x_{l+i} \leftarrow \rightarrow u_i, & i = 1, \ldots, m. \end{array} \right\} \tag{7-42}$$

Hay que recalcar que cualquier transformación de la matriz (7-38) según las reglas del párrafo anterior conduce a una nueva matriz que también va a describir tanto el problema directo como el dual. Por consiguiente, la matriz del tipo más general (7-31) puede servir para describir los problemas de programación lineal tanto directo como dual. Si con esto los elementos de la primera columna (con la excepción posible de α_{00}) son positivos, la matriz corresponde a la solución de base admisible del problema directo. Si son positivos los elementos de la primera fila (con la excepción posible de α_{00}) dicha matriz corresponde a la solución admisible del problema dual. Si en la matriz (7-31) son positivos los elementos tanto de la primera columna como de la fila superior (con la excepción posible de α_{00}), esta matriz corresponde a la solución óptima tanto del problema directo como del dual. Con esto el coeficiente α_{00} brinda el valor de la función de objetivo que, siendo óptima la solución, coincide tanto para el problema directo como para el problema dual:

$$q'_{\text{máx}} = q^*_{\text{mín}}. \tag{7-43}$$

Ejemplo 7-5. El problema de distribución de los recursos está presentado en la tabla 7-3. En las columnas complementarias de esta tabla se da la solución de los problemas directo y dual de programación lineal, es decir, se indica el

plan óptimo de distribución de los recursos, los sobrantes de recursos con el plan óptimo y los precios ocultos. En la tabla se ve que tienen precio nulo los recursos del tipos S_1 que existen en exceso. Tienen el mayor precio los recursos S_2 que se requieren para todos los procesos tecnológicos y cuyas reservas no son grandes.

Tabla 7-3

Datos iniciales para el problema de utilización de los recursos

Tipos de recursos	Gasto de recursos por unidad de producción			Reservas de recursos	Sobrantes de recursos con el plan óptimo	Precios ocultos
	T_1	T_2	T_3			
S_1	2	1	1	25	5,5	0
S_2	1	1	1	14	0	3,0
S_3	0	4	2	19	0	0,5
S_4	3	0	1	24	0	1,0
Ingreso por la realización de la unidad de producción	6	5	5	—	—	—
Plan óptimo	5,5	1,0	7,5	—	—	—

Notemos, en conlcusión, que los precios ocultos pueden jugar un importante papel en diversos casos como instrumento de mando. Así pues, en una gran empresa o rama industrial, en la que muchas decisiones se toman independientemente por los departamentos y grupos de producción, a veces es difícil informar a cada sección sobre las decisiones tomadas por otros departamentos. En este caso los precios ocultos pueden servir de buena orientación a cada departamento para tomar decisiones cercanas a las óptimas desde el punto de vista de los intereses de toda la empresa en conjunto.

d) Concepto de programación de números enteros

En diversos problemas de programación lineal las variables son magnitudes no divisibles en partes. Así, una empresa no puede producir 7,5 aviones, 4,8 turbinas, etc. En estos problemas se añade a las condiciones (7-1) y (7-2) la exigencia de que las variables x_i se expresen por números enteros o, como en el ejemplo 7-4, sean elementos de un conjunto finito. Geométricamente esto significa que las soluciones admisibles serán no todo el dominio de soluciones admisibles definido por las limitaciones (7-1), sino únicamente algunos puntos descritos de este dominio, como se muestra en la figura 7-3.

Desde luego, puede intentarse resolver un problema similar sin tomar en cuenta la condición de que sean enteros los números y hallar la solución determinada en la figura 7-3 por el punto *1*, y después redondear esta solución hasta los números enteros más próximos obteniendo de este modo, en calidad de solución de nú-

meros enteros, el número 2. Sin embargo, con esto puede obtenerse una solución que está muy lejos de ser óptima. Así, en la figura 7-3 el punto *3* va a servir de solución óptima de números enteros.

Existen diversos métodos de solución del problema de programación de números enteros de los cuales el más difundido es el método de Homory. Sin detenernos en el aspecto de cálculo del método de Homory indiquemos solamente la idea general de dicho método.

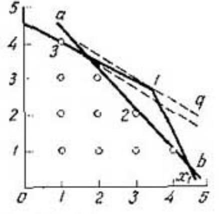


Fig. 7-3. Conjunto de soluciones admisibles en el problema de programación de números enteros

El método de Homory está basado en el empleo del método símplex, con ayuda del cual se busca la solución óptima que no toma en consideración que las variables sean de números enteros o de carácter discreto. Si esta solución resultó de números no enteros, se introduce una limitación adicional que geométricamente representa la línea *ab* en la figura 7-2 la cual secciona una parte del dominio de soluciones admisibles junto con la solución óptima obtenida que no contiene ningún punto admisible de números enteros. Con esta limitación adicional la solución óptima hallada se encuentra fuera del nuevo dominio de soluciones admisibles, es decir, será inadmisible. Por eso el problema se resuelve de nuevo por el método símplex y se halla la solución óptima ya para el nuevo dominio de soluciones admisibles. Si esta nueva solución otra vez no es de números enteros, se repite el proceso.

En los últimos tiempos para resolver el problema de programación de números enteros se eplean con éxito métodos combinatorios entre los cuales el más importante es el método de las ramas y los límites.

PROBLEMAS PARA EL CAPITULO 7

7-1. Transformar las desigualdades (7-10) en ecuaciones. ¿Cuál es el significado físico de las variables añadidas al hacerlo?

7-2. Determinar con ayuda del método símplex el plan óptimo de distribución de los recursos por los datos del ejemplo 7-5, habiendo tomado como variables libres en la etapa inicial x_1, x_2 y x_3.

7-3. Resolver por el método símplex el problema dual por los datos del ejemplo 7-5 habiendo tomado como variables libres en la etapa inicial u_1, u_2, u_5 y u_7. Comparar la matriz correspondiente a la solución óptima con la matriz respectiva, dada en el problema 7-2.

Capítulo octavo
TEORIA DE LOS JUEGOS

8-1. OBJETO DE LA TEORIA DE LOS JUEGOS

a) Juego como modelo de una situación de conflicto

La teoría de los juegos representa una disciplina matemática que se desarrolla intensamente, el objeto de cuyas investigaciones son los métodos de toma de decisión en las llamadas situaciones de conflicto. La situación se llama *de conflicto* si en ella chocan los intereses de varias (habitualmente dos) personas que persiguen objetivos opuestos. Cada una de las partes puede tomar una serie de medidas para alcanzar sus objetivos, con esto(el éxito de una parte significa el fracaso de la otra.

Los autores del primer tratado fundamental sobre la teoría de los juegos, J. von Neumann y O. Morgenstern, se orientaron al análisis de las situaciones de conflicto en las cuestiones de economía cuando, existiendo competencia libre, juegan los papeles de partes en lucha firmas comerciales, empresas industriales, etc. Sin embargo, las situaciones de conflicto se presentan en otras muchas esferas. Se catalogan entre las situaciones de conflicto casi todas las que surgen durante la planificación de operaciones militares, elección del sistema de armamento, protección de los objetos contra ataques, persecución e intercepción de un objetivo, etc. Son ejemplos interesantes de situaciones de conflicto las competiciones deportivas, discusiones de arbitraje, subastas y elecciones al parlamento en caso de haber varios candidatos para un puesto.

Los ejemplos presentados indican la gran variedad de situaciones de conflicto que se encuentran en la práctica. Corrientemente esas situaciones son de difícil análisis directo a cuasa de la multitud de factores secundarios acompañantes. Para hacer posible el análisis matemático de una situación de conflicto es necesario simplificarla, tomando en consideración solamente los factores básicos. El modelo simplificado y formalizado de una situación de conflicto se denomina *juego* y las partes en conflicto (contrincantes), *jugadores.* Nos limitaremos a estudiar los juegos en los que sólo hay partes en conflicto. Antes de dar la definición formal de juego, es necesario aclarar la terminología utilizada que coincide en principo con la que se encuentra en diversos juegos de entretenimiento y azar (ajedrez, damas, juegos de cartas, etc.).

Hay que distinguir el concepto de juego y el de partida individual de este juego. El juego representa la colección de reglas que

describen el comportamiento de los jugadores. Cada caso en que se juega un juego de alguna forma concreta, de principio a fin, constituye una *partida del juego*. Los elementos del juego son las *jugadas*. Las reglas del juego determinan cuál debe ser la secuencia de las jugadas e indican el carácter de cada jugada.

Las jugadas pueden ser personales y aleatorias. La *jugada personal* es la elección por el jugador de una de las variantes de un conjunto preasignado. Por ejemplo, cada jugada de ajedrez es una jugada personal, siendo la primera la elección entre 20 variantes. La decisión tomada por el jugador para la jugada personal se llama *elección*.

La *jugada aleatoria* es también la elección de una de las variantes de un conjunto, pero aquí la variante no se elige por el jugador sino por cierto mecanismo de elección aleatoria. Pueden servir de ejemplos de jugadas aleatorias el reparto de cartas o el lanzamiento de una moneda. La elección realizada durante la jugada aleatoria se denomina *resultado* de esta jugada.

Con respecto a las jugadas las reglas tienen la siguiente estructura.

Para la primera jugada las reglas indican si ésta es una jugada personal o aleatoria. Si es la jugada personal, ellas enumeran las variantes existentes e indican cuál jugador debe hacer la elección. Si es una jugada aleatoria, enumeran las variantes existentes y condicionan las probabilidades de su elección.

Para cada jugada siguiente las reglas determinan según las elecciones y los resultados de las jugadas anteriores:

si va a ser esta jugada personal o aleatoria;

si la jugada va a ser aleatoria, las variantes que se presentan y las probabilidades de su elección;

si la jugada va a ser personal, cuál jugador realiza la elección, las variantes posibles entre las cuales se realiza la elección y la información con respecto a las elecciones y los resultados de las jugadas anteriores.

Por último, las reglas determinan según las elecciones y los resultados de las jugadas que se suceden una tras otra (es decir, en dependencia de la marcha del juego), cuándo debe terminar el juego y cuáles son las ganancias o pérdidas de cada jugador.

b) Concepto de estrategia

Imaginémonos que queremos jugar una partida de ajedrez con las blancas pero no podemos estar presentes personalmente durante el juego. Tenemos un sustituto que debe desarrollar la partida y cumplir todas nuestras indicaciones. Pero él no sabe jugar al ajedrez y no es capaz de tomar decisiones independientes. Para que el sustituto pueda desarrollar toda la partida hasta el final, deben dársele tales indicaciones que prevean cualesquier posiciones posibles en el tablero y determinen para cada posición la jugada

que debe hacerse. El sistema completo de estas indicaciones es la *estrategia*.

Así, la estrategia de las blancas debe indicar la primera jugada, y después para cada respuesta de las negras, la siguiente jugada de las blancas, etc. Por supuesto, la compilación de la estrategia completa durante el juego de ajedrez es un trabajo enorme, prácticamente irrealizable. Por ejemplo, el jugador con las blancas que está presente en persona, debe tomar dos decisiones para hacer dos primeras jugadas. Ahora, jugando con ayuda del sustituto él debe preparar 21 decisiones para aquellas mismas dos jugadas (una decisión, primera jugada y 20 decisiones, respuestas a las 20 primeras jugadas posibles de las negras). No obstante, en muchos problemas más sencillos el concepto de estrategia es sumamente útil.

De esta suerte, la *estrategia* del jugador constituye la descripción unívoca de su elección en cada situación posible en la que él debe hacer una jugada personal.

Si el juego consta solamente de jugadas personales, su resultado está determinado si cada uno de los jugadores eligió su estrategia. Sin embargo, si en el juego hay jugadas aleatorias, entonces él va a tener carácter probabilístico y la elección de la estrategia de los jugadores todavía no determinará definitivamente su resultado.

c) Descripción formal del juego de dos personas

Designemos por X e Y los conjuntos o espacios de todas las estrategias posibles que pueden utilizar los participantes del juego que en lo sucesivo se denominarán, respectivamente, primer y segundo jugadores. Las magnitudes $x \in X$ e $y \in Y$ van a significar las estrategias concretas del primer y segundo jugadores.

Para introducir en el análisis las jugadas aleatorias es cómodo considerar que toma parte en el juego un tercer jugador que hace las jugadas aleatorias utilizando para ello el correspondiente mecanismo de elección aleatoria. Designemos por H el espacio de estrategias de este jugador. Cualquier estrategia $h \in H$ del tercer jugador que representa la secuencia concreta de todas las jugadas aleatorias en la partida ocurrirá con cierta probabilidad $p(h)$ que no es difícil de calcular conociendo las probabilidades de cada jugada aleatoria en esta secuencia. Es fácil ver que $p(h)$ representa la distribución de probabilidades en el espacio H, es decir, satisface las condiciones

$$p(h) \geqslant 0, \quad \sum_{h \in H} p(h) = 1. \qquad (8\text{-}1)$$

Designemos por g cierta variante del juego, es decir, una partida posible. Esta variante estará determinada si se han elegido las estrategias de los jugadores x e y y las de las jugadas aleato-

rias h. Por consiguiente, la partida concreta g representa la tríada de magnitudes x, y, h:

$$g = (x, y, h). \tag{8-2}$$

El resultado de la partida son la ganancia o pérdida de cada uno de los jugadores. Para más comodidad vamos a estimar las ganancias y pérdidas con algún número, por ejemplo, con una suma de dinero en rublos.

Examinemos una de las partidas concretas $g(x, y, h)$ y designemos por $L_x(x, y, h)$ y $L_y(x, y, h)$ las ganancias o pérdidas del primero y segundo jugadores, respectivamente. Al hacerlo, consideramos las ganancias como pérdias negativas. La suma total de pérdidas de ambos jugadores es igual a:

$$L_x(x, y, h) + L_y(x, y, h). \tag{8-3}$$

En lo sucesivo nos limitaremos a examinar solamente los llamados *juegos con suma nula*, es decir, los juegos cuya suma total de pérdidas (8-3) es igual a cero. En tales juegos la pérdida de un jugador es igual a la ganancia del otro.

Al estudiar los juegos con suma nula no hay necesidad de contar por separado las pérdidas o ganancias de ambos jugadores y podemos limitarnos a contar solamente la pérdida del segundo jugador (la ganancia del primer jugador):

$$L_y(x, y, h) = -L_x(x, y, h) = L(x, y, h). \tag{8-4}$$

Por cuanto la estrategia h es aleatoria, con las estrategias elegidas x e y las pérdidas $L(x, y, h)$ serán una magnitud aleatoria con la distribución de probabilidades $p(h)$ en el espacio H. Por eso únicamente pueden estimarse las estrategias elegidas x e y tomando la medida de las pérdidas $L(x, y, h)$ por todo el espacio H, es decir, introduciendo el concepto de pérdidas medias $L(x, y)$ determinadas según (5-64) por la correlación

$$L(x, y) = \sum_{h \in H} L(x, y, h)\, p(h). \tag{8-5}$$

El juego estará determinado si están enumeradas todas las estrategias posibles de los jugadores, es decir, están presentados los espacios X e Y y determinadas las pérdidas $L(x, y)$ para $x \in X$ e $y \in Y$ cualesquiera. De este modo llegamos a la siguiente definición formal de juego.

El juego G se define por la tríada

$$G = (X, Y, L), \tag{8-6}$$

donde X e Y representan cierto espacio y L es una función numérica limitada definida en el producto directo $X \times Y$. Los puntos $x \in X$ e $y \in Y$ se denominan *estrategias* del primero y segundo jugadores y la función L se llama *función de pérdidas.*

Es cómodo presentar los juegos, en los cuales cada jugador tiene un número finito de estrategias (juegos finitos), en forma de la llamada matriz de pérdidas. Sea $G = (X, Y, L)$ un juego finito en el que $X = \{x_1, \ldots, x_m\}$ e $Y = \{y_1, \ldots, y_n\}$. Entonces la matriz de orden $m \times n$

$$\mathbf{Q} = \| q_{ij} \| = \begin{Vmatrix} q_{11} & \cdots & q_{1n} \\ \cdot & \cdot & \cdot \\ q_{m1} & \cdots & q_{mn} \end{Vmatrix}, \qquad (8\text{-}7)$$

en la que $q_{ij} = L(x_i, y_j)$ se denomina *matriz del juego G*.

Para que la descripción del juego esté terminada hay que señalar los objetivos por los cuales se guían los jugadores al elegir sus estrategias. Estos objetivos son lo suficientemente sencillos. El primer jugador trata de asegurarse la mayor ganancia, o sea, de maximizar la función $L(x, y)$ y el segundo jugador trata de hacer mínima su pérdida, es decir, de minimizar la función $L(x, y)$. De este modo, los objetivos de los jugadores resultan directamente opuestos. La dificultad específica con esto es el hecho de que ninguno de los jugadores controla por completo el valor $L(x, y)$ ya que el primer jugador sólo dispone del valor de x, y el segundo, solamente del valor de y. La superación de esta dificultad, o sea, la determinación por cada uno de los jugadores del método más racional para desarrollar el juego, constituye la esencia de la teoría de los juegos.

Es necesario recalcar que los razonamientos expuestos sólo son ciertos para un juego de dos personas con suma nula. Si el número de jugadores es mayor de dos, surge una situación completamente diferente cuya particularidad consiste en que algunos de los jugadores pueden mancomunar sus acciones formando coaliciones con distribución de las ganancias por acuerdo entre los participantes de la coalición. Los participantes de la coalición pueden estimar sus posibilidades con más plenitud y efectuar acciones coordinadas asegurándose así la ganancia más cuantiosa.

Otra variante de juego es el juego con suma no nula. En este juego las ganancias de unos jugadores pueden obtenerse no solamente a cuenta de las ganancias de otros sino también a base de pagos cualesquiera recibidos del exterior. Estos pagos pueden considerarse como pérdidas de cierto jugador adicional ficticio, lo que permite reducir el juego de n personas con suma no nula al juego de $n + 1$ personas con suma nula.

Sin embargo, la teoría de los juegos con n participantes para $n > 2$ es bastante compleja y está insuficientemente investigada. Por eso nos limitaremos solamente al estudio del juego de dos personas con suma nula.

Ejemplo 8-1. Para aclarar los conceptos introducidos, examinemos el juego siguiente que consta de cuatro jugadas.

Primera jugada (personal). El primer jugador elige uno de los dos números enteros 1, 2.

Segunda jugada (aleatoria). Se lanza una moneda y si cae de escudo (y sólo en este caso), se comunica al segundo jugador lo elegido por el primero.

Tercera jugada (personal). El segundo jugador elige uno de los dos números enteros 3, 4.

Cuarta jugada (aleatoria). Se elige de forma aleatoria, con probabilidades 0,4; 0,2; 0,4 uno de los tres números enteros 1, 2, 3.

Resultado del juego: se adicionan los números, elegidos en la primera, tercera y cuarta jugadas y la suma obtenida se paga por el segundo jugador al primero, si ella es par y por el primer jugador al segundo, si es impar.

Durante el análisis preliminar es cómodo representar el juego en forma de árbol en el cual las posiciones que surgen en el proceso del juego se representan por vértices y las jugadas, por ramas que unen un vértice con otro. En la figura 8-1 se muestra el

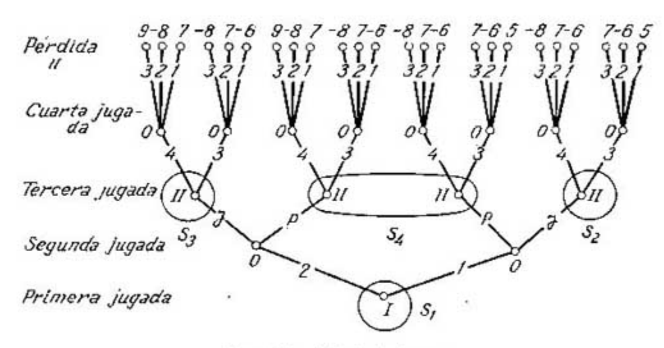


Fig. 8-1 Árbol de juego

árbol del juego considerado. Los vértices correspondientes a las jugadas personales del primero y segundo jugadores, se designan respectivamente por *I* y *II*. Los vértices correspondientes a las jugadas aleatorias se designan por *O*. Los vértices finales que determinan las variantes aisladas del juego están marcadas con cifras que significan las pérdidas del segundo jugador.

En los diferentes vértices correspondientes a las jugadas personales, el jugador posee un tipo determinado de información sobre las jugadas precedentes. Si en varios vértices éste tiene acceso a una misma información, es conveniente reunir estos vértices. Mediante esta reunión se obtienen grupos de vértices S_i denominados *clases de información*. En el ejemplo examinado hay cuatro clases de información cuyo contenido es el siguiente:

S_1, todavía no ha habido jugadas, el primer jugador debe hacer la primera jugada;

S_2, el primer jugador eligió 1;

S_3, el primer jugador eligió 2;

S_4, no se sabe qué eligió el primer jugador.

Al alcanzar al vértice que se halla en la clase S_4, el segundo jugador no tiene información sobre la elección hecha por el primer jugador. Dicho juego se denomina *juego con información incompleta*. Si forma parte de la clase de información sólo un vértice, el jugador que alcanza este vértice está completamente informado sobre todas las jugadas anteriores (él puede obtener esta información, por ejemplo, siguiendo por el árbol del juego el camino hasta el vértice dado), o sea, dispone de toda la información sobre el juego. Los juegos en los que cada clase de información contiene sólo un vértice se denominan *juegos con información completa* (ajedrez, damas, dominó, etc.).

Examinemos el espacio de estrategias de los jugadores. Describamos con la tabla 8-1, *a* el espacio de estrategias del primer jugador que consta sólo de dos elementos, a los cuales corresponde la

Tabla 8-1

Espacio de estrategias para el juego en el ejemplo 8-1

x_1	x_2
(1)	(2)

a)

y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8
(333)	(334)	(343)	(344)	(433)	(434)	(443)	(444)

b)

h	$(J, 1)$	$(J, 2)$	$(J, 3)$	$(P, 1)$	$(P, 2)$	$(P, 3)$
$p(h)$	0,2	0,1	0,2	0,2	0,1	0,2

c)

elección 1 ó 2. La estrategia del segundo jugador debe indicar su jugada con cualquier variante posible de juego. La variante de juego se determina por la clase de información del jugador. Para el segundo jugador hay tres clases de información: S_2, S_3, S_4. Luego, la estrategia del segundo jugador consiste en indicar cuál de los dos número, 3 ó 4, él elige en cada clase de información. Así pues, la estrategia (4, 3, 3) significa que el segundo jugador

elige el 4 en la clase de información S_2 y 3, en las clases de información S_3 y S_4. El espacio de estrategias del segundo jugador se presenta en la tabla 8-1, *b*.

Las estrategias aleatorias h constan de dos jugadas: la segunda es la elección J o P con probabilidades 0,5; 0,5 y la cuarta, la elección del número 1, 2 ó 3 con probabilidades 0,4; 0;2; 0,4. La probabilidad de cada estrategia es igual al producto de las probabilidades de los resultados de ambas jugadas. En la tabla 8-1, *c* se ofrecen el espacio de las estrategias aleatorias H y la distribución de las probabilidades $p(h)$.

Para componer la matriz de pérdidas hay que determinar las pérdidas $q_{ij} = L(x_i, y_j)$ las que en concordancia con (8-5) serán iguales a:

$$q_{ij} = \sum_{h \in H} L(x_i,\ y_j,\ h)\, p(h),$$

con esto las magnitudes $L(x_i, y_j, h)$ representan las cifras con las que están marcados los vértices terminales del árbol del juego.

Es cómodo realizar el cálculo de las magnitudes q_{ij} mediante el método tabular por analogía con la tabla 8-2 preparada para determinar q_{11}.

Tabla 8-2

Cálculo de las pérdidas q_{11} para el juego en el ejemplo 8-1

h	$p(h)$	$L(x_1, y_1, h)$	$L(x_1, y_1, h)\, p(h)$
J, 1	0,2	+5	+1,0
J, 2	0,1	−6	−0,6
J, 3	0,2	+7	+1,4
P, 1	0,2	+5	+1,0
P, 2	0,1	−6	−0,6
P, 3	0,2	+7	+1,4

$q_{11} = +3{,}6$

d) Precios superior e inferior del juego

Para comprender los principios que constituyen el fundamento de la elección por cada jugador de su estrategia, examinemos el juego con la matriz representada en la tabla 8-3.

Supongamos que el primer jugador elige la estrategia x_h. Si la ganancia $L(x_h, y)$ va a depender de la estrategia que elige el segundo jugador, por ejemplo, con la estrategia x_1 las ganancias del primer jugador pueden ser 7, 2, 5, 1. ¿Acaso el primer jugador puede esperar una ganancia máxima igual a 7? Sí, en caso de que éste suponga que el segundo jugador elige la estrategia y_1. No

Tabla 8-3

Matriz del juego con punto silla

	y_1	y_2	y_3	y_4	A (x)
x_1	7	2	5	1	1
x_2	2	2	3	4	2
x_3	5	3	4	4	3*
x_4	3	2	1	6	1
B (y)	7	3*	5	6	

obstante, el segundo jugador puede elegir cualquier otra estrategia, incluso y_4, y entonces la ganancia del primer jugador será igual a 1. Pero ella no puede ser menor de 1 con ninguna estrategia del segundo jugador. Por eso 1, que es el menor elemento del conjunto $L(x_1, y) = \{7, 2, 5, 1\}$, representa la ganancia garantizada del primer jugador con la estrategia x_1.

Generalizando los razonamientos expuestos vemos que si el primer jugador emplea la estrategia x_k, asegura una ganancia garantizada $\mathrm{A}(x_k)$ igual al menor elemento del conjunto $L(x_k, y)$:

$$\mathrm{A}(x_k) = \min_{y} L(x_k, y). \tag{8-8}$$

En la teoría de los juegos se supone que los jugadores obran con suficiente cuidado evitando el riesgo infundado. En este caso el primer jugador debe elegir la estrategia $x \in X$ que corresponde al mayor de los números $\mathrm{A}(x)$. Designando la ganancia garantizada del primer jugador por α y llamándola *precio puro inferior del juego*, obtenemos:

$$\alpha = \max_{x} \mathrm{A}(x) = \max_{x} \min_{y} L(x, y). \tag{8-9}$$

Los valores $\mathrm{A}(x)$, correspondientes al juego con matriz del tipo de la tabla 8-3 se dan en la columna extrema derecha. El valor de α está marcado en esta columna con un asterisco.

Pueden realizarse razonamientos análogos con respecto al segundo jugador. Pero en la matriz del juego se indican sus pérdidas, las que él trata de hacer mínimas. Examinemos la estrategia y_k. Esta estrategia puede proporcionarle una pérdida no mayor de

$$\mathrm{B}(y_k) = \max_{x} L(x, y_k). \tag{8-10}$$

Para asegurarse la pérdida de la menor magnitud, el segundo jugador debe emplear la estrategia $y \in Y$ que corresponde al menor de los números $\mathrm{B}(y)$. Designando la magnitud de la périda,

a la cual puede limitarse el segundo jugador, por β y llamándola *precio puro superior del juego*, obtenemos:

$$\beta = \min_{y} \mathrm{B}(y) = \min_{y} \max_{x} L(x, y). \qquad (8\text{-}11)$$

Los valores B(y) para el juego con matriz del tipo de la tabla 8-3 se muestran en la fila inferior de la tabla. El valor β está marcado en esta línea con un asterisco.

Teorema 8-1. *Si* $G = (X, Y, L)$ *es un juego, para cualquier* $x \in X$ e $y \in Y$ *tiene lugar*:

$$\mathrm{A}(x) \leqslant L(x, y) \leqslant \mathrm{B}(y) \quad \text{y} \quad \alpha \leqslant \beta.$$

Demostración. Por definición,

$$\mathrm{A}(x) = \min_{y} L(x, y) \leqslant L(x, y); \qquad (8\text{-}12)$$

$$\mathrm{B}(y) = \max_{x} L(x, y) \geqslant L(x, y). \qquad (8\text{-}13)$$

Luego,

$$\mathrm{A}(x) \leqslant L(x, y) \leqslant \mathrm{B}(y). \qquad (8\text{-}14)$$

Como que esta correlación se cumple con $x \in X$ e $y \in Y$ cualesquiera, eligiendo en calidad de x el valor para el cual $\mathrm{A}(x) = \alpha$, y a modo de y, el valor para el que $\mathrm{B}(y) = \beta$, obtenemos:

$$\alpha \leqslant \beta. \qquad (8\text{-}15)$$

Como vemos, el precio inferior del juego, es decir, la ganancia que puede asegurársela el primer jugador, no sobrepasa del precio superior del juego, o sea, de la pérdida a que puede limitarse el segundo jugador.

8-2. PRECIOS Y ESTRATEGIAS OPTIMAS DE LOS JUEGOS

a) Juego con punto silla

El caso particular más sencillo del juego es el caso en que $\alpha = \beta$. Designemos esta magnitud por c. Precisamente a este caso pertenece el juego con la matriz presentada en la tabla 8-3, donde $\alpha = \beta = 3$.

En el caso dado ningún método de juego puede garantizar al primer jugador una ganancia mayor de β ya que precisamente el segundo jugador puede limitar su pérdida justamente a la magnitud β. Por otra parte, ningún método de juego puede garantizar al segundo jugador una pérdida menor de α, ya que el primer jugador puede garantizar una ganancia α. Luego, si $\alpha = \beta = c$, ninguna estrategia de ninguno de los jugadores puede asegurarle mejor resultado que c. Al mismo tiempo, cada jugador puede garantizarse el resultado c. En otras palabras, ni el primero ni el

segundo jugadores tienen mejor estrategia que la que les asegura c. En este caso c se denomina *precio puro del juego*, y las estrategias de los jugadores que aseguran el resultado c se llaman *estrategias óptimas*.

La casilla de la matriz que determina la magnitud c se denomina *punto silla*, ya que el valor de c es máximo de la columna y mínimo de la fila en cuya intersección se encuentra esta magnitud. Por eso el juego que tiene precio puro se denomina *juego con punto silla*.

El juego con punto silla se llama *equitativo* si $c = 0$. Si $c \neq 0$, el juego será inequitativo. Para hacerlo equitativo, el primer jugador debe pagarle al segundo la magnitud c antes de comenzar cada nueva partida.

b) Estrategias puras y mixtas

Si el juego no tiene punto silla, surgen dificultades en la determinación de su precio y de las estrategias óptimas de los jugadores. Examinemos, por ejemplo, el juego cuya matriz se da en la tabla 8-4. En este juego $\alpha = 4$ y $\beta = 5$. Por consiguiente, el primer jugador puede garantizarse una ganancia igual a 4 y el segundo puede limitar su ganancia a la magnitud 5. El dominio entre α y β es como si no perteneciera a nadie y cada jugador puede intentar mejorar su resultado a cuenta de este dominio. ¿Cuáles deben ser en este caso las estrategias óptimas de los jugadores?

Tabla 8-4

Matriz del juego sin punto silla

	y_1	y_2	$(A\ x)$
x_1	3	6	3
x_2	5	4	4*
B (y)	5*	6	

Si cada uno de los jugadores emplea la estrategia marcada con el asterisco (x_2 e y_1), la ganancia del primer jugador y la pérdida del segundo van a ser iguales a 5. Esto no es ventajoso para el segundo jugador ya que el primero gana más que los que él puede garantizarse. No obstante, si el segundo jugador descubre de alguna forma el plan del primer jugador de intentar emplear la estrategia x_2, aquél puede utilizar la estrategia y_2 y disminuir la ganancia del primero hasta 4. Es cierto que si el primer jugador adivina la idea del segundo de emplear la estrategia y_2, utilizando la estrategia x_1 el primero aumentará su ganancia hasta 6. De esta forma surge una situación en que cada jugador debe guardar en secreto la estrategia que se prepara a utilizar. Pero, ¿cómo hacerlo? Es que si la partida se juega muchas veces y el segundo jugador emplea todo el tiempo la estrategia y_2, el primer jugador pronto adivinará el pensamiento del segundo y, utilizando la estrategia x_1, tendrá una ganancia adicional. Evidentemente que el se-

gundo jugador debe cambiar de estrategia en cada nueva partida, pero debe hacerlo de manera que el primero no adivine qué estrategia utiliza en cada caso.

Puede guardarse el secreto, si cada vez se elige la estrategia de manera casual, utilizando para ello cualquier mecanismo de elección aleatoria. Por ejemplo, el segundo jugador puede lanzar una moneda y utilizar la estrategia y_1, si cae en escudo, e y_2, si cae en estrella. Tal modo de actuar priva al contrincante de toda posibilidad de conocer por adelantado las acciones de la otra parte.

En caso de utilizarse un mecanismo de elección aleatoria las ganancias y pérdidas de los jugadores serán magnitudes aleatorias. En este caso puede estimarse el resultado del juego por el valor medio de la pérdida del segundo jugador. Así pues, si en el juego con matriz del tipo de la tabla 8-4 el segundo jugador utilizará las estrategias y_1 e y_2 de modo aleatorio con probabilidades 0,5; 0,5, el valor medio de su pérdida con la estrategia del primer jugador x_1 será igual a:

$$L_{med} = q_{11} \cdot 0{,}5 + q_{12} \cdot 0{,}5 = 3 \cdot 0{,}5 + 6 \cdot 0{,}5 = 4{,}5,$$

y con la estrategia del segundo jugador x_2

$$L_{med} = q_{21} \cdot 0{,}5 + q_{22} \cdot 0{,}5 = 5 \cdot 0{,}5 + 4 \cdot 0{,}5 = 4{,}5.$$

Luego, el segundo jugador puede limitar su pérdida media a la magnitud 4,5 independientemente de la estrategia empleada por el primero.

De este modo, en diversos casos resulta racional no trazarse por adelantado la estrategia que debe utilizarse, sino elegir una u otra estrategia de manera aleatoria, fundada en el uso de algún mecanismo de elección aleatoria. Denominemos *estrategia mixta* la basada en la elección aleatoria, a diferencia de las estrategias trazadas por anticipado que examinemos anteriormente y que ahora vamor a llamar *estrategias puras*.

Daremos una definición más rigurosa de las estrategias puras y mixtas.

Sea $G = (X, Y, L)$ un juego. Los espacios $X = \{x_1, \ldots, x_m\}$ e $Y = \{y_1, \ldots, y_n\}$ que contienen la lista de todas las estrategias posibles de los jugadores se denominan *espacios de estrategias puras*, y los elementos de estos espacios $x \in X$ e $y \in Y$ son estrategias puras de los jugadores.

Para obtener una estrategia mixta el jugador debe utilizar cierto mecanismo de elección aleatoria (lanzamiento de una moneda, lanzamiento de un dado de juego, etc.) que tiene un número de resultados igual al número de estrategias puras del jugador.

Supongamos que el mecanismo de elección aleatoria del primer jugador tiene m resultados que constituyen el conjunto $R = \{r^{(1)}, \ldots, r^{(m)}\}$. Designemos por $\xi^{(1)}, \ldots, \xi^{(m)}$ las probabilidades con las que aparecen ciertos resultados del mecanismo de elección aleatoria.

La estrategia mixta del primer jugador consiste en que a cada resultado $r \in R$ se asigna una de las estrategias puras $x \in X$. Con esto las magnitudes $\xi^{(1)}, \ldots, \xi^{(m)}$ van a representar las probabilidades con las cuales se utilizan las estrategias puras $x_1, \ldots, x_m$. El conjunto ordenado $\xi = (\xi^{(1)}, \ldots, \xi^{(m)})$, cuyos elementos satisfacen las condiciones

$$\xi^{(i)} \geqslant 0, \quad \sum_i \xi^{(i)} = 1, \tag{8-16}$$

puede considerarse ahora como la distribución de probabilidades $\xi(x)$ en el espacio X. Esta distribución de probabilidades determina por completo el carácter del juego del primer jugador, y se denomina su *estrategia mixta* correspondiente al mecanismo dado de elección aleatoria.

Otro mecanismo de elección aleatoria proporciona otra distribución de probabilidades $\xi'(x)$, es decir, determina otra estrategia mixta del primer jugador.

En el caso general el primer jugador puede disponer de un número infinito de diferentes mecanismos de elección aleatoria que determinan todas las distribuciones de probabilidades posibles en el espacio de sus estrategias puras:

$$\xi_1 = (\xi_1^{(1)}, \ldots, \xi_1^{(m)}); \quad \xi_2 = (\xi_2^{(1)}, \ldots, \xi_2^{(m)}) \ldots \tag{8-17}$$

Entonces, el conjunto

$$\mathrm{E} = \{\xi_1, \xi_2, \ldots\} \tag{8-18}$$

va a representar *el espacio de las estrategias mixtas del primer jugador.*

De suerte análoga a ésta el segundo jugador puede utilizar su mecanismo de elección aleatoria, que determina las probabilidades $\eta^{(1)}, \ldots, \eta^{(n)}$ con las que van a utilizarse las estrategias puras $y_1, \ldots, y_n$. En este caso el conjunto ordenado $\eta = (\eta^{(1)}, \ldots, \eta^{(n)})$, cuyos elementos satisfacen las correlaciones

$$\eta^{(k)} \geqslant 0, \quad \sum_k \eta^{(k)} = 1, \tag{8-19}$$

representa la distribución de probabilidades $\eta(y)$ en el espacio Y y se llama *estrategia mixta del segundo jugador.*

El segundo jugador, al igual que el primero, puede disponer de un número infinito de diferentes mecanismos de elección aleatoria, que determinan diversas distribuciones de probabilidades en el espacio de sus estrategias puras:

$$\eta_1 = (\eta_1^{(1)}, \ldots, \eta_1^{(n)}; \quad \eta_2 = (\eta_2^1, \ldots, \eta_2^{(n)}), \ldots, \tag{8-20}$$

el total de las cuales

$$\mathrm{H} = \{\eta_1, \eta_2, \ldots\} \tag{8-21}$$

constituye *el espacio de estrategias mixtas del segundo jugador.*

c) Función de pérdidas al utilizar estrategias mixtas

Examinemos el juego $G = (X, Y, L)$, en el que $X = \{x_1, \ldots, x_m\}$ e $Y = \{y_1, \ldots, y_n\}$ son los espacios de estrategias puras de los jugadores y $L(x, y)$ representa las pérdidas del segundo jugador determinadas para las estrategias puras $x \in X$ e $y \in Y$. En lo sucesivo vamos a denominar la función $L(x, y)$ *función de pérdidas.*

Supongamos ahora que los jugadores emplean estrategias mixtas. Esto significa que están presentados todos los conjuntos $\Xi = \{\xi_1, \xi_2, \ldots\}$ y $H = \{\eta_1, \eta_2, \ldots\}$, cuyos elementos representan las estrategias mixtas de los jugadores, es decir, las diversas distribuciones de probabilidades $\xi(x)$ y $\eta(y)$ en los espacios X e Y. El carácter del juego estará definido, si cada jugador eligió su estrategia mixta $\xi \in \Xi$ y $\eta \in H$. Por consiguiente, en caso de utilizarse estrategias mixtas, deben formar parte de la definición del juego los conjuntos Ξ y H en lugar de los conjuntos X e Y.

Al utilizar las estrategias mixtas cambia asimismo la magnitud de la pérdida del segundo jugador, o sea, la función de pérdidas. Por cuanto el juego adquiere carácter aleatorio, serán aleatorias las magnitudes de las ganancias y pérdidas de los jugadores. Por eso ahora puede hablarse únicamente de la magnitud media de la ganancia X o la pérdida Y que se definirá tanto por la función de pérdidas $L(x, y)$ como por la distribución de probabilidades $\xi(x)$ y $\eta(y)$ y puede hallarse por la fórmula para el valor medio de la función de dos variables como

$$L(\xi, \eta) = \sum_{x, y} L(x, y)\, \xi(x)\, \eta(y). \tag{8-22}$$

La función de pérdidas $L(\xi, \eta)$ debe formar parte de la definición del juego en caso de utilizarse las estrategias mixtas.

De este modo, en caso de utilizarse las estrategias mixtas obtenemos en lugar del juego $G = (X, Y, L)$ un nuevo juego $\Gamma = (\Xi, H, L)$ que se llama promedio del juego G. En este juego la función de pérdidas $L(\xi, \eta)$ se determina por la fórmula (8-22) como la media de $L(x, y)$ con las distribuciones de probabilidades presentadas $\xi(x)$ y $\eta(y)$.

Todo lo dicho más arriba puede compendiarse en forma de las siguientes definiciones.

Sea $G = (X, Y, L)$ un juego. Supongamos que Ξ y H determinan diferentes distribuciones de probabilidades ξ y η en los espacios X e Y, y $L(\xi, \eta)$ representa el valor medio de la función de pérdidas. Entonces el juego

$$\Gamma = (\Xi, H, L) \tag{8-23}$$

se denomina *promedio* del juego G.

Sea $\Gamma = (\Xi, H, L)$ el promedio del juego $G = (X, Y, L)$. Entonces los puntos $x \in X$ e $y \in Y$ se llaman *estrategias puras* en el

juego G, y los puntos $\xi \in \Xi$ y $\eta \in H$ se denominan *estrategias mixtas* en el juego G.

Subrayamos que las estrategias puras pueden considerarse como un caso particular de estrategias mixtas. Efectivamente, la estrategia mixta determinada por la distribución de probabilidades

$$\xi(x) = \begin{cases} 1 & \text{para} \quad x = x_k; \\ 0 & \text{para} \quad x \neq x_k \end{cases} \tag{8-24}$$

coincide con la estrategia pura x_k y la estrategia mixta determinada por la distribución de probabilidades

$$\eta(y) = \begin{cases} 1 & \text{para} \quad y = y_l; \\ 0 & \text{para} \quad y \neq y_l \end{cases} \tag{8-25}$$

coincide con la estrategia pura y_l.

d) Precios superior e inferior del juego al utilizar estrategias mixtas

Sean $G = (X, Y, L)$ el juego y $\Gamma = (\Xi, H, L)$ el promedio del juego.

Supongamos que el primer jugador emplea la estrategia mixta $\xi \in \Xi$. A este jugador le interesa cuál va a ser su ganancia garantizada, es decir, la suma mínima de ganancia que puede asegurarse con certeza, incluso si el segundo jugador utiliza su mejor estrategia mixta η. Con esto no se excluye el caso posible de que el segundo jugador emplee alguna de las estrategias puras.

Designemos por $A(\xi)$ la magnitud de ganancia asegurada del primer jugador con la estrategia ξ. Evidentemente, esto es el límite inferior de la función de ganancia $L(\xi, \eta)$ con la ξ dada y diferentes $\eta \in H$, es decir,

$$A_\Gamma(\xi) = \min_\eta L(\xi, \eta). \tag{8-26}$$

Más adelante, al primer jugador le va a interesar elegir entre todas las estrategias posibles $\xi \in \Xi$ aquella con la que su ganancia garantizada será máxima, es decir, le va a interesar el límite superior de la función $A_\Gamma(\xi)$ designado por α_Γ, y denominada *precio inferior del juego* con las estrategias mixtas de los jugadores:

$$\alpha_\Gamma = \max_\xi A_\Gamma(\xi) = \max_\xi \min_\eta L(\xi, \eta). \tag{8-27}$$

La estrategia para la cual se cumple la condición (8-27) se denomina estrategia *de minimax* (a veces de maximin) del primer jugador y se designa por ξ_0.

Supongamos ahora, que el segundo jugador eligió alguna estrategia mixta $\eta \in H$. A este jugador le interesa, cuál será su pérdida máxima con la mayor estrategia $\xi \in \Xi$ del primer jugador, o sea,

el límite superior de la función $L(\xi, \eta)$ con η dada y distintas $\xi \in \mathrm{E}$, el que se designa por $\mathrm{B}_\Gamma \eta)$:

$$\mathrm{B}_\Gamma(\eta) = \max_{\xi} L(\xi, \eta). \tag{8-28}$$

Al segundo jugador le va a interesar la elección de tal estrategia $\eta \in \mathrm{H}$ con la que su pérdida será mínima, es decir, le va a interesar el límite inferior de la función $\mathrm{B}_\Gamma(\eta)$ designado por β_Γ y llamado *precio superior del juego* con estrategias mixtas de los jugadores:

$$\beta_\Gamma = \min_{\eta} \mathrm{B}_\Gamma(\eta) = \min_{\eta} \max_{\xi} L(\xi, \eta). \tag{8-29}$$

La estrategia, para la que se cumple la condición (8-29), se llama estrategia *de minimax* del segundo jugador y se designa por η_0.

Para facilitar los razonamientos ulteriores introduzcamos las siguientes designaciones para el precio inferior y superior del juego $G = (X, Y, L)$ al utilizar estrategias puras:

$\mathrm{A}_G(x) = \min\limits_{y} L(x, y)$ es la ganancia garantizada del primer jugador con la estrategia $x \in X$;

$\alpha_G = \max\limits_{x} \mathrm{A}_G(x)$ es el precio inferior del juego G con estrategias puras de los jugadores;

$\mathrm{B}_G(y) = \max\limits_{x} L(x, y)$ es la pérdida mayor del segundo jugador con la estrategia $y \in Y$;

$\beta_G = \min\limits_{y} \mathrm{B}_G(y)$ es el precio superior del juego G con estrategias puras de los jugadores.

Teorema 8-2. *si se tiene el juego $G = (X, Y, L)$ y $\Gamma = (\mathrm{E}, \mathrm{H}, L)$ es el promedio del juego G, entonces*

$$\left.\begin{aligned}
&\text{a)}\ \mathrm{A}_\Gamma(\xi) = \min_{y} L(\xi, y);\\
&\text{b)}\ \alpha_G \leqslant \alpha_\Gamma;\\
&\text{c)}\ \mathrm{B}_\Gamma(\eta) = \max_{x} L(x, \eta);\\
&\text{d)}\ \beta_G \geqslant \beta_\Gamma;\\
&\text{e)}\ \alpha_\Gamma \leqslant \beta_\Gamma.
\end{aligned}\right\} \tag{8-30}$$

El punto a) contiene la afirmación de que, al escoger el primer jugador cualquier estrategia mixta, la magnitud de su ganancia garantizada es igual a la magnitud de la ganancia garantizada al utilizar el segundo jugador solamente estrategias puras. Según el punto b) el precio inferior del juego con estrategias mixtas del primer jugador no es menor que el precio inferior del juego con estrategias puras, o sea, existe una estrategia mixta que en todo caso no es peor que la estrategia pura óptima. Los puntos c) y

d) contienen afirmaciones análogas respecto al segundo jugador. El punto e) significa que el precio inferior del juego α_Γ al utilizar estrategias mixtas no excede del precio superior del juego, es decir, al emplear las mejores estrategias mixtas la ganancia garantizada del primer jugador no sobrepasa de la pérdida asegurada del segundo jugador.

Demostración. a) Como que las estrategias puras $y \in Y$ son un caso particular de las mixtas $\eta \in$ E, entonces $L(\xi, y)$ forma parte como caso particular de $L(\xi, \eta)$, es decir, constituye un subconjunto de $L(\xi, \eta)$. Basándonos en el teorema de límite inferior del subconjunto (véase el § 1-1), tenemos:

$$\min_y L(\xi, y) \geqslant \min_\eta L(\xi, \eta) = \mathrm{A}_\Gamma(\xi). \tag{8-31}$$

Más adelante $L(\xi, \eta)$ puede considerarse como la media de $L(\xi, y)$, tomada por todas las y con la distribución de probabilidades $\eta(y)$. Basándonos en el teorema del valor medio (véase el § 5-5), obtenemos:

$$\min_y L(\xi, y) \leqslant L(\xi, \eta). \tag{8-32}$$

La correlación (8-32) es válida para cualquier $\eta \in$ H, inclusive con η_0, para la cual $L(\xi, \eta_0) = \min\limits_\eta L(\xi, \eta)$. De esta manera,

$$\min_y L(\xi, y) \leqslant \min_\eta L(\xi, \eta) = \mathrm{A}_\Gamma(\xi). \tag{8-33}$$

Comparando (8-31) y (8-33), nos convenceremos de la certeza del punto a).

b) Comparando $\mathrm{A}_G(x) = \min\limits_y L(x, y)$ con la magnitud $\mathrm{A}_\Gamma(\xi)$, la cual, fundándose en el punto a), puede definirse también como $\mathrm{A}_G(x)$ sólo con estrategias puras y del segundo jugador y teniendo en cuenta que la estrategia pura x es un caso particular de la estrategia mixta $\xi(x)$, vemos $\mathrm{A}_G(x)$ es el subconjunto $\mathrm{A}_\Gamma(\xi)$. Basándonos en el teorema de límite superior del subconjunto, tenemos:

$$\max_x \mathrm{A}_G(x) \leqslant \max_\xi \mathrm{A}_\Gamma(\xi), \tag{8-34}$$

lo que demuestra el punto b).

Los puntos c) y d) se demuestran análogamente.

e) Por definición tenemos:

$$\left.\begin{aligned} \mathrm{A}_\Gamma(\xi) &= \min_\eta L(\xi, \eta) \leqslant L(\xi, \eta), \\ \mathrm{B}_\Gamma(\eta) &= \max_\xi L(\xi, \eta) \geqslant L(\xi, \eta). \end{aligned}\right\} \tag{8-35}$$

Por consiguiente,

$$\mathrm{A}_\Gamma(\xi) \leqslant L(\xi, \eta) \leqslant \mathrm{B}_\Gamma(\eta). \tag{8-36}$$

Esto es válido para ξ y η cualesquiera, es decir, para una ξ con la cual $A_\Gamma(\xi) = \alpha_\Gamma$, y para una η con la que $B_\Gamma(\eta) = \beta_\Gamma$, lo que demuestra el punto e).

Corolario. Mediante la comparación de los puntos b), d) y e) hallamos:

$$\alpha_G \leqslant \alpha_\Gamma \leqslant \beta_\Gamma \leqslant \beta_G. \tag{8-37}$$

Fe (8-37) se infiere que si el juego G tiene precio puro, es decir, $\alpha_G = \beta_G = c$ (como vimos, esto tiene lugar para los juegos con punto silla), entonces $\alpha_\Gamma = \beta_\Gamma = c$, o sea, el juego Γ también tiene precio puro. Por eso la estrategia óptima en el juego G es asimismo óptima en el juego Γ. En este caso el juego Γ no debe considerarse en absoluto y las estrategias óptimas de los jugadores pueden hallarse por el método descrito para los juegos con punto silla.

Ahora bien, el juego Γ fue introducido con el objetivo de analizar aquellos juegos en los que $\alpha_G < \beta_G$. Surge la pregunta, no resulta acaso para los juegos de gran clase, que en este caso $\alpha_\Gamma = \beta_\Gamma = c_\Gamma$ y la magnitud c_Γ puede considerarse como el precio puro del juego Γ. A continuación demostraremos que esto ocurre realmente así: el empleo de estrategias mixtas permite hallar el precio del juego y las estrategias óptimas de los jugadores también en este caso. Al hacerlo, vamos a partir de la siguiente definición de estrategia óptima.

Supongamos que existe el juego $G = (X, Y, L)$ y sea $\Gamma = (\Xi, H, L)$ el promedio del juego G. Si el juego Γ tiene el precio puro $c_\Gamma = \alpha_\Gamma = \beta_\Gamma$, se dice, que G tiene el precio c_Γ y cualquier estrategia óptima en Γ, es decir, la estrategia que asegure valores garantizados de la ganancia del primer jugador y de la pérdida del segundo, iguales a c_Γ, se denomina estrategia *óptima* en el juego G. No es difícil ver que si $\alpha_\Gamma = \beta_\Gamma = c_\Gamma$, las estrategias máximas de los jugadores $\xi_0(x)$ y $\eta_0(y)$ serán estrategias óptimas.

El teorema fundamental de la teoría de los juegos es el teorema referente a que cada juego finito tiene precio y cada uno de los jugadores tiene estrategias óptimas. El mismo será demostrado en el párrafo siguiente.

8-3. TEOREMA FUNDAMENTAL DE LA TEORIA DE LOS JUEGOS

a) Juego *S*

Puede darse una interpretación geométrica útil a los juegos en los cuales el primer jugador tiene un número finito de estrategias puras. Supongamos que está dado el juego $G = (X, Y, L)$ con la matriz $\mathbf{Q} = \| q_{ij} \|$, definida por la expresión (8-7). Relacionemos con cada $y \in Y$ el punto C en el espacio m-dimensional

cuyas coordenadas son las perdidas del segundo jugador con todas las estrategias puras posibles del primer jugador:

$$C = [L(x_1, y), \ldots, L(x_m, y)]. \qquad (8\text{-}38)$$

Tomando en consideración que $L(x_i, y_j) = q_{ij}$, y dando a y todos los valores posibles, obtendremos n puntos:

$$C_1 = (q_{11}, \ldots, q_{m1}), \ldots; \quad C_n = (q_{1n}, \ldots, q_{mn}), \qquad (8\text{-}39)$$

que tienen en calidad de sus coordenadas las pérdidas dadas en las columnas correspondientes de la matriz de juego. El juego presentado por el conjunto de puntos

$$\{C_1, \ldots, C_n\}, \qquad (8\text{-}40)$$

recibió el nombre de *juego S*. Se juega a éste del modo siguiente.

El segundo jugador elige uno de los puntos C_i del conjunto (8-40), por ejemplo C_3. Independientemente de la elección del segundo jugador, el primer jugador elige una coordenada del punto C_i, por ejemplo la segunda. Con esto, las pérdidas del segundo jugador van a ser iguales al valor de la segunda coordenada del punto C_3, es decir, q_{23}.

Como vemos, el juego S es equivalente a un juego corriente, por lo menos, al utilizar estrategias puras, ya que la elección de un punto del conjunto (8-40) es equivalente a la elección de la estrategia $y \in Y$, y la de la coordenada de este punto, equivalente a la elección de la estrategia $x \in X$.

Si el número de estrategias del primer jugador es igual a dos, el juego S tiene una clara representación geométrica ya que los puntos del conjunto (8-40) serán en este caso puntos del plano.

Ejemplo 8-2. Examinemos el juego con la matriz definida por la tabla 8-5. El juego S equivalente contiene cinco puntos: $C_1 = (1, 0)$, $C_2 = (2, 3)$, $C_3 = (-1, 1)$, $C_4 = (0, -1)$, $C_5 = (1, 2)$. La representación geométrica de este juego se muestra en la figura 8-2.

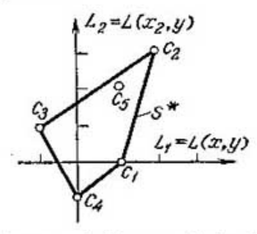


Fig. 8-2. Representación geométrica del juego S

Designemos por S^* la envoltura convexa del conjunto finito de puntos $\{C_1, \ldots, C_n\}$ y demostremos el teorema, del cual se infiere que el juego S es equivalente a un juego corriente al utilizar estrategias no sólo puras sino mixtas.

Tabla 8-5

	y_1	y_2	y_3	y_4	y_5
x_1	1	2	—1	0	1
x_2	0	3	1	—1	2

Teorema 8-3. *Cualquier estrategia mixta del segundo jugador puede representarse por un punto perteneciente a la envoltura convexa S^* y, a la inversa, cualquier punto $S \in S^*$ puede considerarse como cierta estrategia mixta del segundo jugador.*

Demostracion. Examinemos las estrategias mixtas de los jugadores $\xi = (\xi^{(1)}, \ldots, \xi^{(m)})$ y $\eta = (\eta^{(1)}, \ldots, \eta^{(n)})$. Las pérdidas del segundo jugador, al utilizar los datos de las estrategias mixtas, serán iguales a:

$$L(\xi, \eta) = \sum_i \sum_k q_{ik}\xi^{(i)}\eta^{(k)} = \sum_i \xi^{(i)} \sum_k q_{ik}\eta^{(k)} = \sum_i \xi^{(i)}S^{(i)}, \quad (8\text{-}41)$$

donde

$$S^{(i)} = \sum_k q_{ik}\eta^{(k)}. \quad (8\text{-}42)$$

Designemos por S el punto en el espacio m-dimensional con las coordenadas $S^{(1)}, \ldots, S^{(m)}$, en el que, según (8-42),

$$\left.\begin{array}{l} S^{(1)} = q_{11}\eta^{(1)} + \ldots + q_{1n}\eta^{(n)}; \\ \ldots\ldots\ldots\ldots\ldots \\ S^{(m)} = q_{m1}\eta^{(1)} + \ldots + q_{mn}\eta^{(n)}. \end{array}\right\} \quad (8\text{-}43)$$

Teniendo en cuenta que $(q_{1i}, \ldots, q_{mi}) = C_i$, podemos escribir la expresión (8-43) en forma de correlación vectorial

$$S = C_1\eta^{(1)} + \ldots + C_n\eta^{(n)} = \sum_i C_i\eta^{(i)}. \quad (8\text{-}44)$$

Tomando en consideración que las magnitudes $\eta^{(i)}$ satisfacen las correlaciones (8-19), vemos que S no es otra cosa que la media ponderada de los puntos $C_1, \ldots, C_n$ con pesos $\eta^{(1)}, \ldots, \eta^{(n)}$ y, consecuentemente, S es cierto punto perteneciente a la envoltura convexa S^*. De esta manera, a cada estrategia $\eta = (\eta^{(1)}, \ldots, \eta^{(n)})$ del segundo jugador va a corresponder cierto punto perteneciente a la envoltura convexa S^* y la presentación de este punto es equivalente a la presentación de la estrategia mixta del segundo jugador.

Se cumple asimismo lo inverso. Por cuanto cualquier punto S, perteneciente a la envoltura convexa S^* puede representarse como la media ponderada de los puntos $C_1, \ldots, C_n$, que determinan la envoltura convexa S^*, o sea, está representada en forma de la expresión (8-44), para cada punto $S \in S^*$ se hallarán pesos $\eta^{(1)}, \ldots$ $\ldots, \eta^{(n)}$ cuya presentación determinará la estrategia mixta del segundo jugador.

Corolario. Por cuanto la estrategia mixta del primer jugador en el juego S se mantiene igual que en un juego corriente, del teorema demostrado se deduce que el juego S es completamente equivalente al juego corriente, es decir, cualquier juego puede representarse en forma del juego S equivalente.

En lo sucesivo vamos a designar por Γ_1 el juego S. Para pasar del juego $\Gamma(\mathrm{E}, \mathrm{H}, L)$ al S, en lugar del espacio de estrategias mixtas del segundo jugador $\mathrm{H} = \{\eta_1, \eta_2, \ldots\}$ es necesario utilizar el espacio de estrategias S, o sea, la envoltura convexa S^*. Si se designan por L_1 las pérdidas del segundo jugador en el juego S, dicho juego se define por la tríada E, S^* y L, por cierto, las pérdidas L_1 deben determinarse por el producto directo $\mathrm{E} \times S^*$. De esta suerte, la expresión

$$\Gamma_1 = (\mathrm{E}, S^*, L_1) \tag{8-45}$$

define el juego S. La función de pérdidas $L_1(\xi, S)$ se determina con esto fundándose en (8-41) de la manera siguiente:

$$L_1(\xi, S) = \sum_i \xi^{(i)} S^{(i)} = \xi S, \tag{8-46}$$

donde por ξS se designa el producto escalar de los vectores $\xi = (\xi^{(1)}, \ldots, \xi^{(m)})$ y $S = (S^{(1)}, \ldots, S^{(m)})$.

b) Precios inferior y superior del juego en el juego S

Si el primer jugador emplea en el juego S la estrategia mixta $\xi \in \mathrm{E}$, la magnitud de su ganancia sea igual a:

$$\mathrm{A}_{\Gamma_1}(\xi) = \min_S L_1(\xi, S) = \min_S \xi S. \tag{8-47}$$

Designemos por $\xi_0 = (\xi_0^{(1)}, \ldots, \xi_0^{(m)})$ la estrategia del primer jugador con la que $\mathrm{A}_{\Gamma_1}(\xi)$ alcanza el máximo. El valor de $\mathrm{A}_{\Gamma_1}(\xi)$ será igual al precio inferior del juego α_{Γ_1} el cual debido a la equivalencia del juego S con el ordinario coincide con α_Γ. Por lo tanto,

$$\alpha_\Gamma = \mathrm{A}_{\Gamma_1}(\xi_0) = \max_\xi \mathrm{A}_{\Gamma_1}(\xi) = \max_\xi \min_S \xi S. \tag{8-48}$$

La estrategia ξ_0 que satisface la correlación (8-48) se denomina estrategia *minimax* del primer jugador.

Supongamos ahora que el segundo jugador emplea cierta estrategia $S \in S^*$. La magnitud de la pérdida, a la cual él se limita con esto, será igual a:

$$\mathrm{B}_{\Gamma_1}(S) = \max_\xi L_1(\xi, S) = \max_\xi \xi S. \tag{8-49}$$

Designemos por $S_0 = (S_0^{(1)}, \ldots, S_0^{(m)})$ un punto tal de la envoltura convexa S^*, que minimiza la magnitud $\mathrm{B}_{\Gamma_1}(S)$. El valor $\mathrm{B}_{\Gamma_1}(S_0)$ va a ser igual al precio superior del juego β_{Γ_1}, el que

debido a la equivalencia del juego S con el ordinario coincide con β_Γ. Así pues,

$$\beta_\Gamma = B_{\Gamma_1}(S_0) = \min_S B_{\Gamma_1}(S) = \min_S \max_\xi \xi S. \qquad (8\text{-}50)$$

La estrategia S_0 que satisface la correlación (8-50) se denomina estrategia *minimax* del segundo jugador.

Las expresiones para $B_{\Gamma_1}(S)$ y β_Γ puede presentarse en forma más cómoda si utilizamos el teorema siguiente.

Teorema 8-4. *Si S es un punto arbitrario del espacio m-dimensional y $\xi = (\xi^{(1)}, \ldots, \xi^{(m)})$, la variable multidimensional que satisface la condición* (8-16), *se cumple la correlación*

$$\max_\xi \xi S = \max(S^{(1)}, \ldots, S^{(m)}). \qquad (8\text{-}51)$$

Demostración. Sea $S^{(k)} = \max(S^{(1)}, \ldots, S^{(m)})$. Examinemos el valor particular ξ, correspondiente al caso

$$\xi^{(i)} = \begin{cases} 1 & \text{para} \quad i = k; \\ 0 & \text{para} \quad i \neq k. \end{cases} \qquad (8\text{-}52)$$

En este caso $\xi S = S^{(k)}$. De esta suerte, $S^{(k)}$ es el valor particular del producto escalar ξS y, por consiguiente, el subconjunto del conjunto de valores ξS que se obtienen con todos los valores posibles de ξ. Basándonos en el teorema de límite superior del subconjunto, hallamos:

$$S^{(k)} = \max(S^{(1)}, \ldots, S^{(m)}) \leqslant \max_\xi \xi S. \qquad (8\text{-}53)$$

Por otra parte, sustituyendo en la expresión para ξS las magnitudes $S^{(1)}, \ldots, S^{(m)}$ por su valor máximo $S^{(k)}$, obtenemos

$$\xi S = \sum_i \xi^{(i)} S^{(i)} \leqslant S^{(k)} \sum_i \xi^{(i)} = S^{(k)}. \qquad (8\text{-}54)$$

Esta expresión es válida con cualquier ξ, que satisfaga la correlación (8-16), inclusive con dicha ξ cuando ξS se convierte en máximo. Comparando (8-53) y (8-54) llegamos a la correlación (8-51).

Utilizando el teorema demostrado, representemos la expresión para $B_{\Gamma_1}(S)$ en la forma

$$B_{\Gamma_1}(S) = \max_\xi \xi S = \max(S^{(1)}, \ldots, S^{(m)}). \qquad (8\text{-}55)$$

Deduzcamos varios corolarios de la expresión (8-55).

Corolario 1. De la condición $\beta_\Gamma \leqslant B_{\Gamma_1}(S)$ obtenemos:

$$\beta_\Gamma \leqslant \max(S^{(1)}, \ldots, S^{(m)}). \qquad (8\text{-}56)$$

Cualquier punto $S \in S^*$ tiene, por lo menos, una coordenada no menor de β_Γ (o sea, mayor o igual).

Corolario 2. Suponiendo que $S = S_0 = (S_0^{(1)}, \ldots, S_0^{(m)})$, obtenemos:

$$\beta_\Gamma = B_{\Gamma_1}(S_0) = \text{máx}\,(S_0^{(1)}, \ldots, S_0^{(m)}). \tag{8-57}$$

El precio superior del juego β_Γ es igual al máximo de las coordenadas del punto S_0 que determina la estrategia minimax del segundo jugador.

c) Teorema de minimax

La posibilidad de que cada jugador encuentre su mejor estrategia se funda en el siguiente teorema, que puede considerarse como demostración de la existencia de solución para los juegos finitos.

Teorema 8-5. *Cualquier juego finito tiene precio y cada jugador tiene por lo menos una estrategia óptima.*

Premisas de partida. Sean $G = (X, Y, L)$ un juego finito y $\Gamma = (\text{E}, \text{H}, L)$, el promedio de este juego. Durante la demostración del teorema será cómodo razonar en términos del juego S. Por eso, designaremos por $\Gamma_1 = (\text{E}, S^*, L_1)$ el juego S equivalente.

Los precios inferior y superior del juego van a ser respectivamente iguales a α_Γ y β_Γ independientemente de si examinamos el juego Γ o el juego Γ_1 equivalente al juego S, siendo $\alpha_\Gamma \leqslant \beta_\Gamma$ como se mostró anteriormente.

Hara demostrar el teorema es suficiente mostrar que $\beta_\Gamma \leqslant \alpha_\Gamma$ ya que de la comparación con la desigualdad anterior se va a inferir que $\alpha_\Gamma = \beta_\Gamma$, es decir, que el juego tiene precio. Para demostrar esta desigualdad es suficiente hallar tal estrategia mixta ξ_0 del primer jugador, con la que para todas las $S \in S^*$ tiene lugar:

$$\beta_\Gamma \leqslant \xi_0 S, \quad S \in S^*. \tag{8-58}$$

En efecto, si tiene lugar (8-58), entonces

$$\beta_\Gamma \leqslant \min_S \xi_0 S = A_\Gamma(\xi_0) = \alpha_\Gamma. \tag{8-59}$$

De este modo, la demostración del teorema se reducirá a demostrar la desigualdad (8-58).

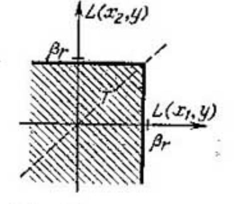


Fig. 8-3. Dominio T en el espacio bidimensional

Demostración. Examinemos el conjunto T que consta de todos los puntos $t = (t^{(1)}, \ldots, t^{(m)})$, tales que $t^{(i)} < \beta_\Gamma$ con $i = 0, 1, \ldots, m$. En la figura 8-3 se muestra el dominio T para el espacio bidimensional que tiene en el caso dado el aspecto de cuña rectangular con vértice situado en la recta trazada desde el origen de coordenadas a un ángulo de 45° con el eje de las abscisas. Aclaremos algunas propiedades del conjunto T.

El conjunto T es convexo. Examinemos dos puntos arbitrarios $\mathbf{t}_1$ y t_2 de este conjunto. La ecuación del segmento que une estos dos puntos va a tener la forma:

$$t = w_1 t_1 + w_2 t_2; \quad w_1, \quad w_2 \geqslant 0; \quad w_1 + w_2 = 1. \tag{8-60}$$

Proyectando esta ecuación en el eje i y tomando en cuenta el teorema 8-4, obtenemos:

$$t^{(i)} = w_1 t_1^{(i)} + w_2 t_2^{(i)} \leqslant \text{máx}\,(t_1^{(i)},\ t_2^{(i)}) < \beta_\Gamma. \tag{8-61}$$

Por consiguiente, cualquier punto del segmento considerado pertenece a T y el conjunto T es convexo.

El conjunto T no se interseca con el conjunto S^*. Esto se desprende de que cualquier punto del conjunto S^* tiene, por lo menos, una coordenada mayor o igual a β_Γ (véase el corolario 1 del teorema 8-4) lo que significa que T y S^* no tienen puntos comunes.

Por cuanto T y S^* son dominios convexos no intersecados, existe un hiperplano que los divide de tal suerte que los conjuntos T y S^* resultarán situados en los diferentes semiespacios determinados por este hiperplano. Por lo tanto, existen cierta $a = ({}^{(1)}, \ldots$ $\ldots, a^{(m)})$ y un número c tales que la ecuación

$$ax = c \tag{8-62}$$

será la ecuación del hiperplano divisor, siendo

$$\left.\begin{aligned} aS &\geqslant c \quad \text{para} \quad S \in S^*; \\ at &\leqslant c \quad \text{para} \quad t \in T. \end{aligned}\right\} \tag{8-63}$$

Demostremos que $a^{(i)} \geqslant 0$, $i = 1, \ldots, m$. Sea $\delta_i = (\delta_i^{(1)}, \ldots$ $\ldots, \delta_i^{(m)})$, un punto en el que la coordenada i es igual a 1, y restantes son iguales a la pequeña magnitud $\varepsilon > 0$. Examinemos el punto $S_0 \in S^*$. Por cuanto su coordenada máxima es igual a β_Γ (véase el corolario 2 del teorema 8-4), el punto $S_0 - \delta_i \in T$. Por consiguiente,

$$aS_0 \geqslant c \geqslant a\,(S_0 - \delta_i). \tag{8-64}$$

De aquí se deduce que

$$a\delta_i = a^{(1)}\delta_i^{(1)} + \ldots + a^{(m)}\delta_i^{(m)} \geqslant 0. \tag{8-65}$$

Si $\varepsilon \to 0$, entonces $\delta_i^{(k)} \to 0$ con $k \neq i$ y $\delta_i^{(i)} = 1$. Con esto, la última condición da:

$$a^{(i)} \geqslant 0, \quad i = 1, \ldots, m. \tag{8-66}$$

Introduzcamos la designación

$$\xi_0 = \frac{a}{\sum_i a^{(i)}}. \tag{8-67}$$

Evidentemente, $\xi_0 \in E$, ya que

$$\xi_0^{(i)} = \frac{a^{(i)}}{\sum_i a^{(i)}} \geqslant 0, \quad \sum_i \xi_0^{(i)} = 1. \qquad (8\text{-}68)$$

Introduzcamos, además, la designación

$$\nu = \frac{c}{\sum_i a^{(i)}}. \qquad (8\text{-}69)$$

Dividamos las desigualdades (8-63) por $\sum_i a^{(i)}$. Tomando en cuenta (8-67) y (8-68), obtenemos:

$$\left.\begin{array}{l} \xi_0 S \geqslant \nu \quad \text{para} \quad S \in S^*; \\ \xi_0 t \leqslant \nu \quad \text{para} \quad t \in T. \end{array}\right\} \qquad (8\text{-}70)$$

Examinemos el punto t_0 con las coordenadas $t_0^{(i)} = \beta_\Gamma - \varepsilon$, $\varepsilon > 0$, $i = 1, \ldots, m$. Evidentemente $t_0 \in T$. Basándonos en la segunda de las designaciones (8-70), obtenemos:

$$\xi_0 t_0 \leqslant \nu. \qquad (8\text{-}71)$$

Supongamos que $\varepsilon \to 0$, de modo que $t_0^{(i)} \to \beta_\Gamma$. Entonces

$$\xi_0 t_0 = \sum_i \xi_0^{(i)} t_0^{(i)} \to \beta_\Gamma \sum_i \xi_0^{(i)} = \beta_\Gamma. \qquad (8\text{-}72)$$

Comparando (8-71) y (8-72), hallamos que

$$\nu \geqslant \beta_\Gamma. \qquad (8\text{-}73)$$

Con esto, la primera de las desigualdades (8-70) da:

$$\xi_0 S \geqslant \nu \geqslant \beta_\Gamma, \qquad (8\text{-}74)$$

lo que demuestra la desigualdad (8-58).

De esta manera $\nu = \alpha_\Gamma = \beta_\Gamma$ es el precio del juego, constituyendo ξ_0 y S_0 las estrategias mixtas óptimas de los jugadores.

d) Representación geométrica del principio de minimax

Mostremos que el punto S_0 que determina la estrategia de minimax del segundo jugador, es un punto límite del dominio S^* tal, en que el dominio T es tangente al S^*.

Como vimos, $S_0 = (S_0^{(1)}, \ldots, S_0^{(m)}) \in S^*$, con esto mäx $\times$ $\times (S_0^{(1)}, \ldots, S_0^{(m)}) = \beta_\Gamma$. Por otro lado, $t = (t^{(1)}, \ldots, t^{(m)} \in T$, si $t^{(i)} < \beta_\Gamma$, $i = 1, \ldots, m$.

Examinemos el punto $S_0' = (S_0^{(1)} - \varepsilon, \ldots, S_0^{(m)} - \varepsilon)$, donde $\varepsilon > 0$. Evidentemente, $\max(S_0^{(1)} - \varepsilon, \ldots, S_0^{(m)} - \varepsilon) < \beta_\Gamma$, así que $S_0^{(i)} - \varepsilon < \beta_\Gamma$, $i = 1, \ldots, m$. Luego, $S_0' \in T$. Pero el $\lim\limits_{\varepsilon \to 0} S_0' = S_0 \in S^*$.

De aquí se derivan dos deducciones:

1) S_0 es un punto límite del dominio S^*;

2) S_0 es un punto en el cual el dominio T es tangente al dominio S^*.

Estas propiedades permiten hallar geométricamente con facilidad la estrategia de minimax S_0 para el caso cuando el primer jugador tiene dos estrategias puras, es decir, cuando el juego S equivalente se representa mediante el conjunto de puntos del plano. Para construir el dominio T, tangente al dominio S^*, es

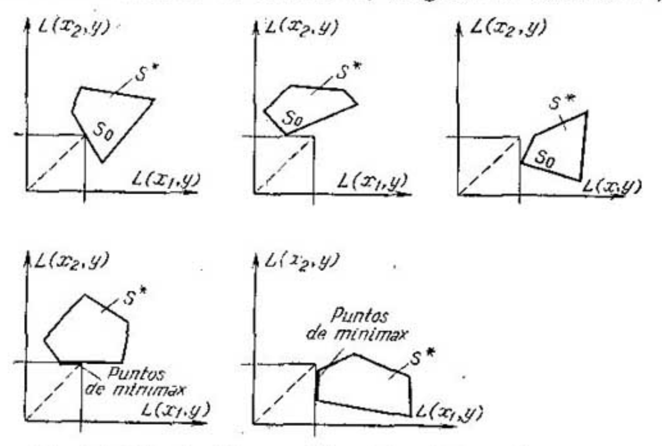


Fig. 8-4. Determinación geométrica de la estrategia de minimax

cómodo trazar una recta auxiliar del origen de coordenadas, a un ángulo de 45°, al eje de las abscisas en el cual se encuentra el vértice de la cuña rectangular que forma el dominio T.

En la figura 8-4 se muestran los diferentes casos de situación mutua de los dominios S^* y T y están marcados los puntos que determinan la estrategia de minimax S_0 del segundo jugador.

8-4. RESOLUCION DE LOS JUEGOS

a) Estrategias predominantes y útiles

Como sabemos, las estrategias mixtas de los jugadores construyen una mezcla de estrategias puras $x \in X$ e $y \in Y$, que se toman en correspondencia con la distribución de las probabilidades $\xi(x)$ y $\eta(y)$. No obstante, en muchos casos el empleo de algunas de las estrategias puras es manifiestamente irracional y al determinar la estrategia mixta óptima las mismas simplemente no de-

ben tomarse en consideración. Vamos a denominar *estrategias útiles* del jugador aquellas de las estrategias puras que forman parte de la estrategia mixta óptima. Para distinguir con mayor facilidad las estrategias útiles, introduzcamos el concepto de estrategias predominantes.

Examinemos dos estrategias y_l e y_p del segundo jugador. Supongamos que el primero emplea la estrategia x_i. Las pérdidas del segundo serán, respectivamente, q_{il} y q_{ip}. Puede resultar que con cualquier i

$$q_{il} \leqslant q_{ip}, \quad i = 1, \ldots, m, \tag{8-75}$$

es decir, en la matriz de juego las pérdidas en la columna l no superan las pérdidas correspondientes en la columna p. Esto significa, que al segundo jugador bajo ningunas condiciones le es ven-

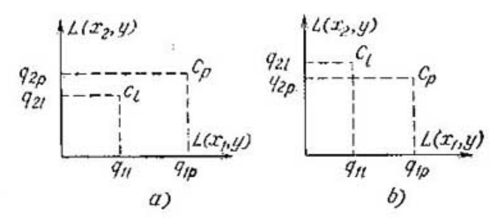


Fig. 8-5. Determinación del predominio en el juego S

tajoso emplear la estrategia y_p, ya que empleándola, incurre a sabiendas en pérdidas mayores que con la estrategia y_l. Por eso la estrategia y_p deberá ser omitida, es decir, borrada de la matriz de juego. La estrategia y_l que satisface la condición (8-75) se denomina *predominante* con respecto a la estrategia y_p.

Las estrategias predominantes del segundo jugador tienen clara representación geométrica al pasar al juego S equivalente en el plano. En este caso $m = 2$ y la condición (8-75) se escribirá en la forma

$$q_{1l} \leqslant q_{1p}, \quad q_{2l} \leqslant q_{2p}. \tag{8-76}$$

En la figura 8-5 se muestran dos casos de disposición de los puntos C_l y C_p correspondientes a las estrategias puras y_l e y_p del segundo jugador. Es fácil ver que en el caso representado en la figura 8-5, *a* la estrategia C_l predomina sobre la C_p, y en el representado en la figura 8-5, *b* ninguna de las estrategias es predominante. Para que la estrategia y_l predomine sobre la y_p, el punto C_l debe encontrarse más a la izquierda y más abajo del punto C_p.

De manera análoga se determinan las estrategias predominantes del primer jugador. Decimos que la estrategia x_l predomina sobre la x_p si la ganancia del primer jugador con la estrategia x_l es

mayor que las ganancias con la estrategia x_p con cualquier estrategia $y \in Y$:

$$q_{li} \geqslant q_{pi}, \quad i = 1, \ldots, n, \tag{8-77}$$

es decir, si en la matriz de juego las pérdidas en la fila x_l son mayores que las pérdidas correspondientes de la fila x_p.

Ejemplo 8-3. En la figura 8-6 el juego S se presenta en el plano con los puntos $C_1 - C_6$. Como se ve en la figura, C_1 predomina sobre el punto C_5 y C_1, C_2 y C_6 predominan sobre el C_4. Eliminando los puntos C_5 y C_4, llegamos al juego S determinado por los puntos C_1, C_2, C_3, C_6 entre los cuales no hay predominantes.

La eliminación de la matriz de juego, de aquellas estrategias sobre las cuales predominan otras, simplifica considerablemente el juego y, por consiguiente, también la búsqueda de la estrategia óptima. Supongamos ahora que en el juego considerado ninguna estrategia predomina sobre otra. Cabe preguntar si son útiles todas las estrategias en este caso, es decir, si se utilizan para obtener la estrategia mixta óptima.

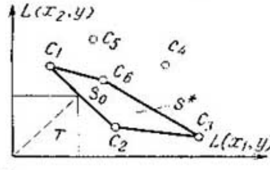


Fig. 8-6. Ejemplo de juego S

Sirvámonos del juego presentado en la figura 8-6. Después de eliminar los puntos C_4 y C_5 sobre los cuales predominan otros puntos, llegamos al juego representado por los puntos C_1, C_2, C_3 y C_6. Una vez construido el dominio T, vemos que el punto S_0, que determina la estrategia óptima del segundo jugador se encuentra en la recta que une los puntos C_1 y C_2 y puede representarse como la media ponderada de dichos puntos. De este modo, la estrategia mixta óptima del segundo jugador constituye la mezcla de las estrategias puras y_1 e y_2, que son útiles en el caso dado. Las estrategias y_3 e y_6 no son útiles y no es racional emplearlas.

El número total de estrategias útiles de cada uno de los jugadores puede determinarse partiendo de que la envoltura convexa S^* del conjunto finito $\{C_1, \ldots, C_n\}$ es un poliedro convexo en el espacio m-dimensional. El punto S_0, que es un punto límite del poliedro convexo S^*, va a pertenecer sin falta a una de sus caras cuyos vértices corresponderán a las estrategias últiles del segundo jugador. Tomando en consideración que el número de vártices de cualquier cara del poliedro convexo S^* no puede exceder del número total de sus vértices, o sea el número n, y no puede exceder de la dimensionalidad del espacio en el cual existe el poliedro convexo, o sea, del número m, llegamos a la conclusión de que el número de estrategias del segundo jugador no sobrepasa del menor de los números m y n. Por cuanto las nociones del primero y del segundo jugadores son convencionales, también es cierta la deducción análoga para el primer jugador.

De esta suerte, en el juego con matriz de tamaño $m \times n$ el número de estrategias útiles de cada uno de los jugadores no sobrepasa del menor de los números m y n.

Teorema 8-6. *Si uno de los jugadores se atiene a su estrategia mixta óptima, la ganancia de los jugadores se mantiene invariable e igual al precio del juego* ν, *independientemente de cuál estrategia mixta (o pura) emplea el otro jugador, si éste no sale de los límites de sus estrategias útiles.*

Demostración. Sea ν el precio del juego y $\xi_0=(\xi_0^{(1)}, \ldots, \xi_0^{(m)})$ la estrategia mixta óptima del primer jugador. Si el primer jugador emplea su estrategia óptima ξ_0, su ganancia no puede ser menor de ν.

Supongamos que el segundo jugador tiene k estrategias útiles que se designan por $y_1, \ldots, y_k$. Si el primer jugador emplea su estrategia óptima ξ_0 y el segundo utiliza la estrategia útil $y_i (i = 1, \ldots, k)$, la ganancia del primer jugador ν_i no será menor de ν:

$$\nu_i \geqslant \nu, \quad i=1, \ldots, k. \tag{8-78}$$

Examinemos ahora la estrategia mixta óptima del segundo jugador η_0. Esta se presenta por medio de las probabilidades $\eta_0^{(i)}$ con las que se emplean las estrategias puras útiles. Lo esencial es que no puede ser $\eta_0^{(i)}=0$ con $i = 1, \ldots, k$, ya que en este caso, la i-éstima estrategia no formaría parte de la estrategia mixta óptima del segundo jugador, es decir, no sería útil.

Hallemos la ganancia media del primer jugador si el segundo emplea su estrategia mixta óptima. El empleo de la estrategia mixta óptima η_0 por el segundo jugador, significa que el primer jugador va a obtener las ganancias $\nu_1, \ldots, \nu_k$ correspondientes a las estrategias puras $y_1, \ldots, y_k$ con las probabilidades $\eta_0^{(1)}, \ldots$ $\ldots, \eta_0^{(k)}$. Entonces, la ganancia media del primer jugador va a ser igual a:

$$L_{\text{med}}=\nu_1\eta_0^{(1)}+\ \ldots\ +\nu_k\eta_0^{(k)}=\sum_{i=1}^{k}\nu_i\eta_0^{(i)}. \tag{8-79}$$

Sustituyendo ν_i por ν y teniendo en cuenta (8-78), obtenemos:

$$L_{\text{med}}\leqslant \nu\sum_{i=1}^{k}\eta_0^{(i)}=\nu. \tag{8-80}$$

Pero la ganancia media con las estrategias óptimas de los jugadores es el precio del juego ν, es decir:

$$L_{\text{med}}=\nu. \tag{8-81}$$

Por otro lado, la correlación (8-79) puede pasar a ser (8-81) únicamente, si se cumple la condición

$$\nu_i=\nu, \quad i=1, \ldots, k. \tag{8-82}$$

lo que demuestra el teorema.

De este modo, la ganancia media del primer jugador, si éste emplea su estrategia óptima mixta, será la misma en el caso de que el segundo jugador emplee cualquiera de las estrategias útiles, es decir, cualquier estrategia mixta que conste de estrategias útiles.

Claro está que debe recordarse que siempre lo más ventajoso es utilizar la estrategia mixta óptima. El empleo de las estrategias útiles aunque no óptimas no acarrea daños sólo en el caso de que el contrincante se atenga a la estrategia óptima. La utilización de la estrategia óptima siempre asegura una ganancia igual al precio del juego.

Ahora queda por esclarecer, cómo se hallan las estrategias útiles y óptimas de los jugadores, es decir, estudiar la cuestión de la búsqueda de la resolución del juego.

b) Búsqueda de las estrategias óptimas

Examinemos el juego $G=(X,\ Y,\ L)$ con la matriz $m\times n$. Vamos a considerar que en el juego G falta el punto silla (en caso contrario la solución se halla fácilmente por la regla establecida en el § 8-2) y se han eliminado las estrategias sobre las cuales existe predominio.

Designemos por $\xi_0=(\xi_0^{(1)}, \ldots, \xi_0^{(m)})$ la estrategia mixta óptima del primer jugador. Algunas $\xi_0^{(i)}$ pueden ser iguales a cero. Esto significa que la estrategia correspondiente x_i no es útil. Ante nosotros se plantea el problema de hallar los valores de $\xi_0^{(i)}$ para $i=1, \ldots, m$.

Supongamos que el segundo jugador empleará la estrategia y_k. Si dicha estrategia es útil, la ganancia del primer jugador va a ser igual a v, pero si la estrategia no es útil, la ganancia del primer jugador puede ser hasta mayor de v. Luego, en el caso general se cumple que

$$L(\xi_0,\ y_k)=\xi_0^{(1)}q_{1k}+\ \ldots\ +\xi_0^{(m)}q_{km}\geqslant v. \qquad (8\text{-}83)$$

Tales expresiones pueden formarse para cada estrategia pura del segundo jugador, es decir, para $i=1, \ldots, n$. Además, se sabe que

$$\xi_0^{(i)}\geqslant 0, \quad \xi_0^{(1)}+\ \ldots\ +\xi_0^{(m)}=1. \qquad (8\text{-}84)$$

Vamos a considerar que $v>0$. Esto va a ocurrir siempre, si todos los elementos de la matriz de juego q_{ik} son positivos. Si entre los elementos q_{ik} hay los negativos, éstos pueden hacerse positivos adicionando a todos los elementos de la matriz de juego cierto número $v'>0$. Entonces el precio del juego también aumenta en la magnitud v',

Dividamos las ecuaciones (8-83) y (8.84) por ν, designando

$$\frac{\xi_0^{(i)}}{\nu}=p_i. \qquad (8\text{-}85)$$

Es evidente que $p_i \geqslant 0$. Las desigualdades (8-83) con $k = 1, \ldots, n$ se escribirán entonces en la forma:

$$\left.\begin{array}{l} p_1q_{11}+\ \ldots\ +p_m a_{m^1}\geqslant 1; \\ p_1q_{1n}+\ \ldots\ +p_m q_{mn}\geqslant 1, \end{array}\right\} \qquad (8\text{-}86)$$

y la correlación (8-84) tomará la forma:

$$p_1+\ \ldots\ +p_m=\frac{1}{\nu}. \qquad (8\text{-}87)$$

Forman parte de las desigualdades lineales (8-86) las magnitudes incógnitas $p_1, \ldots, p^m$ que determinan la estrategia mixta del primer jugador. La determinación de la estrategia mixta óptima requiere la búsqueda de una solución del sistema (8-86) tal que la magnitud ν se vuelve máxima y, por consiguiente, la forma lineal (8-87) adquiere el valor mínimo. Pero esto es un problema corriente de programación lineal.

Para más comodidad se debe pasar en las expresiones (8-86) de desigualdades a igualdades, substrayendo de los segundos miembros las incógnitas positivas adicionales z_i, lo que da el sistema de las ecuaciones

$$p_1q_{1i}+\ \ldots +\ p_mq_{mi}-z_i=1,\ i=1,\ \ldots,\ n. \qquad (8\text{-}88)$$

En las ecuaciones (8-88) que corresponden a las estrategias útiles del segundo jugador, como resultado de la solución se obtendrá que $z_i = 0$. De este modo, resolviendo el sistema (8-88) bajo la condición de que sea minimizada la forma lineal (8-87), hallamos la estrategia mixta del primer jugador, el precio del juego ν y las estrategias útiles del segundo jugador $y_1, \ldots, y_k$.

Para buscar la estrategia mixta óptima del segundo jugador $\eta_0=(\eta_0^{(1)}, \ldots, \eta_0^{(k)})$, que contiene k incógnitas $\eta_0^{(1)}, \ldots, \eta_0^{(k)}$, es necesario formar k ecuaciones. Una de ellas es

$$\eta_0^{(1)}+\ \ldots\ +\eta_0^{(k)}=1. \qquad (8\text{-}89)$$

Obtenemos las $k-1$ ecuaciones restantes formulando expresiones para las pérdidas medias con la estrategia mixta óptima del segundo jugador y para $k-1$ estrategias útiles cualesquiera del primer jugador. Así, para la estrategia útil x_i vamos a tener la ecuación

$$\eta_0^{(1)}q_{i1}+\ \ldots\ +\eta_0^{(k)}q_{ik}=\nu. \qquad (8\text{-}90)$$

Al componer $k-1$ ecuaciones de la forma (8-90) y al adicionarles la ecuación (8-89), obtendremos k ecuaciones resolviendo

las cuales es fácil hallar las magnitudes $\eta_0^{(1)}, \ldots, \eta_0^{(k)}$, que determinan la estrategia mixta óptima η_0 del segundo jugador.

Ejemplo 8-4. Hallar la solución del juego con la matriz presentada en la tabla 8-6.

Tabla 8-6

	y_1	y_2	y_3
x_1	2	—3	4
x_2	—3	4	—5
x_3	4	—5	6

Tabla 8-7

	y_1	y_2	y_3
x_1	7	2	9
x_2	2	9	0
x_3	9	0	11

Ante todo cerciorémonos de que en el juego falta el punto silla y ninguna de las estrategias predomina sobre las otras. Para no tener que operar con elementos negativos de la matriz del juego sumamos a cada elemento de la matriz el número 5. La matriz del juego adquiere el aspecto de la tabla 8-7.

Las ecuaciones (8-88) se escriben en la forma

$$7p_1 + 2p_2 + 9p_3 - z_1 = 1;$$
$$2p_1 + 9p_2 - z_2 = 1;$$
$$9p_1 + 11p_3 - z_3 = 1.$$

La solución de estas ecuaciones debe satisfacer la condición de mínimo de la forma lineal (8-87)

$$p_1 + p_2 + p_3 = \text{mín.}$$

Resolviendo el problema de programación lineal obtenido, hallamos:

$$z_1 = z_2 = z_3 = 0; \quad p_1 = 0{,}05; \quad p_2 = 0{,}1; \quad p_3 = 0{,}05.$$

Como vemos, todas las tres estrategias y_1, y_2 e y_3 son útiles. De la expresión (8-87) encontramos:

$$\nu = \frac{1}{p_1 + p_2 + p_3} = 5.$$

Con esto

$$\xi_0^{(1)} = 5p_1 = 0{,}25; \quad \xi_0^{(2)} = 5p_2 = 0{,}5; \quad \xi_0^{(3)} = 5p_3 = 0{,}25.$$

Para la búsqueda de la estrategia mixta óptima del segundo jugador formamos una ecuación del tipo (8-89)

$$\eta_0^{(1)} + \eta_0^{(2)} + \eta_0^{(3)} = 1$$

y dos ecuaciones del tipo (8-90) para las estrategias útiles del primer jugador x_2 y x_3

$$\left.\begin{aligned} 2\eta_0^{(1)} + \ \ 9\eta_0^{(2)} &= 5; \\ 9\eta_0^{(1)} + 11\eta_0^{(3)} &= 5. \end{aligned}\right\}$$

Resolviendo conjuntamente las tres ecuaciones obtenidas, encontramos:

$$\eta_0^{(1)} = \eta_0^{(3)} = 0{,}25; \quad \eta_0^{(2)} = 0{,}5.$$

Recordando que se adicionó el número 5 a todos los elementos de la matriz, hallamos el precio del juego $\nu - 5 = 0$.

c) Representación geométrica del principio de minimax en el juego de $2 \times n$

En el juego de $2 \times n$ la estrategia mixta del primer jugador representa el par ordenado $\xi = (\zeta, 1 - \zeta)$. Si el segundo jugador emplea la estrategia y_k, la ganancia media del primero es

$$L(\zeta, y_k) = \zeta q_{1k} + (1 - \zeta) q_{2k} = q_{2k} + \zeta (q_{1k} - q_{2k}), \qquad (8\text{-}91)$$

o sea, depende linealmente de la magnitud ζ como se muestra en la fig. 8-7.

La determinación de la estrategia óptima ξ_0 se examinará en el ejemplo del juego con la matriz presentada en la tabla 8-8. En la figura 8-8 están designadas con las cifras *1*, *2* las líneas de ganancia para las estrategias y_1, y_2 e y_3, respectivamente. La ganancia garantizada del primer jugador con cualquier ζ va a determinarse por el límite inferior de la gráfica dada, marcada con línea gruesa. Según el principio de minimax, la estrategia óptima ζ_0 será aquella con la que es máxima la ganancia garantizada. Dicha estrategia se define por el punto M determinado por la intersección de las líneas de ganancias con las estrategias y_1 e y_2 que serán de este modo estrategias útiles del segundo jugador.

Tabla 8-8

	y_1	y_2	y_3
x_1	1	3	5
x_2	4	2	1

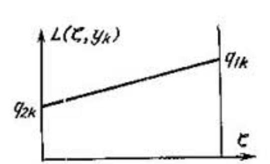


Fig. 8-7. Líneas de ganancia con la estrategia y_k

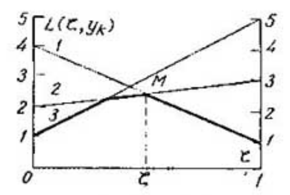


Fig. 8-8. Estrategia de minimax en el juego $2 \times m$

La estrategia mixta óptima ξ_0 puede hallarse por las condiciones de igualdad de la ganancia media del primer jugador con las estrategias útiles y_1 e y_2 del segundo:

$$\zeta_0 + 4(1 - \zeta_0) = 3\zeta_0 + 2(1 - \zeta_0),$$

lo que da $\zeta_0 = 1/2$. De este modo, $\xi_0 = (1/2, 1/2)$. El precio del juego se halla ahora como la ganancia media del primer jugador con la estrategia ξ_0 y cualquiera de las estrategias útiles, por ejemplo, y_1 del segundo:

$$\nu = \zeta_0 + 4(1 - \zeta_0) = 2{,}5.$$

La estrategia mixta óptima del segundo jugador $\eta_j = (\eta_0^{(1)}, 1 - \eta_0^{(1)}, 0)$ con el precio conocido v del juego se encuentra por la expresión de las pérdidas medias del segundo jugador con cualquiera de las estrategias útiles del primero. Por elemplo, con la estrategia x_1 obtenemos:

$$\eta_j^{(1)} + 3\left(1 - \eta_j^{(1)}\right) = 2{,}5,$$

lo que da $\eta_j^{(1)} = {}^1/_4$. Así pues, $\eta_0 = ({}^1/_4, {}^3/_4, 0)$.

PROBLEMAS PARA EL CAPITULO 8

8-1. Calcule la matriz de pérdidas para el juego dado en el ejemplo 8-1.

8-2. Calcule la matriz de pérdidas para el juego siguiente dibujando previamente el árbol del juego.

Se lanza una moneda simétrica y sí cae en escudo, el primer jugador elige uno de los números 1, 6. Si cae en estrella éste elige uno de los números 2, 7. Después, el segundo jugador elige uno de los dos números 3, 9. Los números elegidos por los jugadores se adicionan y dan la suma S. Después se hace el sorteo con una probabilidad de 0,8 de caída en escudo. Si en este caso la moneda cae en escudo, el segundo jugador paga al primero S rublos, si cae en estrella, el primer jugador paga al segundo S rublos. El primer jugador conoce el resultado de la primera jugada aleatoria. El segundo no sabe nada acerca de las jugadas anteriores.

8-3. Un cómodo necanismo de elección aleatoria es la posición del segundero en un cuadrante de reloj en un momento aleatorio de tiempo. Describa el empleo de este mecanismo de elección aleatoria para obtener las distribuciones de probabilidades siguientes:

$$\left(\frac{3}{5}, \frac{2}{5}\right), \quad \left(\frac{1}{6}, \frac{2}{3}, \frac{1}{6}\right), \quad \left(\frac{1}{5}, \frac{2}{5}, \frac{2}{5}\right).$$

8-4. Decpriba las acciones de los jugadores en una partida del juego cuya matriz está presentada en la tabla 8-4, en caso de utilizar el primer jugador la estrategia pura x_2, y el segundo, la estrategia mixta $\eta = (0{,}5; 0{,}5)$.

Capítulo noveno

TEORIA DE LAS DECISIONES ESTADISTICAS (JUEGOS ESTADISTICOS)

9-1. ESTRUCTURA DE LOS JUEGOS ESTADISTICOS

a) Juegos estratégicos y estadísticos

Los llamados *juegos estadísticos* son un tipo específico de juegos que tienen gran importancia al analizar distintas situaciones prácticas. Ellos tienen una serie de diferencias substanciales con aquel tipo de juegos que se estudiaron hasta ahora y que pueden ser llamados *juegos estratégicos.*

La base de la teoría de los juegos estratégicos es la suposición de que los intereses de los dos jugadores son opuestos. Cada uno de los jugadores trata de elegir su estrategia de tal suerte que obtenga la mayor ventaja para sí y reduzca al mínimo la ventaja del adversario. En tales juegos cada jugador obra activamente y trata de utilizar en lo posible su estrategia óptima.

Sin embargo, en muchas situaciones de la práctica nos encontramos con casos en que uno de los jugadores resulta *neutral,* es decir, no trata de sacar para sí la ventaja máxima y, por tanto, no tiende a revertir en su beneficio los errores cometidos por el contrario. Se catalogan entre dichos juegos aquellos en los que la naturaleza obra en calidad de uno de los jugadores. Aquí entendemos por el término "la naturaleza", el total de circunstancias exteriores en cuyas condiciones es necesario tomar la decisión.

La naturaleza no puede considerarse como un contrario racional que pudiera utilizar los errores cometidos por el hombre. En otras palabras, la naturaleza no tiene mala intención con respecto al hombre. Ella simplemente se desarrolla y obra en correspondencia con sus leyes y está en manos del hombre revertir estas leyes en su beneficio. Si el hombre conociera con perfecta exactitud las leyes de la naturaleza, pudiera utilizarlas con beneficio máximo para sí. Pero en muchos casos el hombre no conoce la ley de la naturaleza o no la conoce suficientemente bien.

El pago ineludible por el intento de obtener una decisión en condiciones de información incompleta sobre la ley de la naturaleza es la posibilidad de tomar decisiones erróneas. Con esto, en la práctica las situaciones pueden ser tales que hacen imposible renunciar por completo a tomar alguna decisión. Además, la decisión de renunciar a tomar una decisión es también una decisión y puede tener consecuencias tan indeseables como otras decisiones.

La única salida de la situación creada es la elaboración por el hombre de tal estrategia respecto a la toma de decisión, que aunque no excluya la posibilidad de tomar decisiones incorrectas, sin embargo, reduzca al mínimo las consecuencias indeseables motivadas.

Es cierto que el hombre tiene además la posibilidad de estudiar a su adversario, es decir, a la naturaleza, mediante la realización de un experimento.

Teóricamente, realizando un experimento infinito, podemos hacer tan completos como querramos nuestros conocimientos de la naturaleza obrando ya en condicioens de completa determinación. No obstante, dos circunstancias lo dificultan:

para efectuar el experimento se necesita tiempo mientras que en muchos casos hay que tomar la decisión rápidamente;

el experimento requiere gastos materiales, es decir, puede ser costoso, más costoso que la ganancia que proporcionarán los conocimientos adicionales obtenidos en el experimento.

Por eso un importante problema del hombre en su juego contra "la naturaleza" es tomar la decisión de si hace falta realizar el experimento, y si hace falta, cuál, cuándo se terminará y qué operaciones se realizarán después de terminado el mismo.

Los juegos en los cuales un adversario es la naturaleza y el otro, el hombre, recibieron la denominación de *juegos estadísticos*, y la teoría de dichos juegos se denomina *teoría de las decisiones estadísticas*. En lo sucesivo vamos a llamar *estadista* al hombre que toma parte en el juego contra la naturaleza.

b) Espacio de estrategias de la naturaleza

Vamos a entender por estrategia de la naturaleza el total de condiciones exteriores en las cuales resulta necesario tomar una decisión. Llamaremos estado de la naturaleza ϑ a este total de condiciones exteriores. En el caso general existe cierto conjunto de estados posibles de la naturaleza $\Theta = \{\vartheta_1, \ldots, \vartheta_m\}$ que, como acordamos en el capítulo 6, vamos a llamar espacio de estados de la naturaleza. Vamos a denominar a los elementos ϑ_i de este espacio estrategias puras de la naturaleza.

Si supiéramos por anticipado cuál de sus estrategias puras emplea la naturaleza en cada caso concreto, utilizaríamos con seguridad la decisión basada en el conocimiento completo del estado de la naturaleza. Sin embargo, por lo común sólo se conoce la lista de estrategias puras de la naturaleza. Además, de la experiencia del pasado se conoce cuán frecuentemente la naturaleza emplea una u otra de sus estrategias puras, es decir, se conoce la distribución apriorística de probabilidades $\xi(\vartheta)$ en el espacio de estados de la naturaleza Θ. Vamos a denominar *estrategia mixta de la naturaleza* esta distribución apriorística de las probabilidades $\xi(\vartheta)$.

c) Espacio de estrategias del estadista y función de pérdidas

La tarea del estadista consiste en tomar alguna decisión o realizar alguna operación del total de decisiones u operaciones. Designemos por $a_1, \ldots, a_l$ las operaciones posibles del estadista. Cada una de ellas es una *estrategia pura* del estadista. El conjunto $A = \{a_1, \ldots, a_l\}$ es el *espacio de estrategias puras del estadista.*

El estadista debe saber estimar cada una de sus operaciones. Para ello éste admite que llevando a cabo la operación a, puede sufrir una pérdida $L(\vartheta, a)$, que depende tanto de la operación ejecutada a como del estado desconocido de la naturaleza ϑ. La función $L(\vartheta, a)$ denominada *función de pérdidas*, debe estar determinada anticipadamente para todas las combinaciones posibles de $a \in A$ y $\vartheta \in \Theta$, es decir, debe estar presentada con el producto directo de los conjuntos $\Theta \times A$. Ella puede presentarse analíticamente, o por analogía con (8-7), mediante una matriz de pérdidas de la forma

$$\mathbf{Q} = | q_{ij} | = \begin{Vmatrix} q_{11} & \ldots & q_{1l} \\ \cdot & \cdot \; \cdot & \cdot \\ q_{m1} & \ldots & q_{ml} \end{Vmatrix}, \qquad (9\text{-}1)$$

donde $q_{ij} = L(\vartheta_i, a_j)$. El conocimiento de la función de pérdidas permite al estadista efectuar las operaciones que son mejores en las condiciones de información sobre el estado de la naturaleza de que dispone.

Por lo común el estadista conoce la estrategia mixta de la naturaleza, o sea, la distribución apriorística de las probabilidades $\xi(\vartheta)$ en el espacio de estados de la naturaleza Θ. El conocimiento de la distribución apriorística de las probabilidades permite determinar las pérdidas medias que sufre el estadista realizando una u otra operación:

$$L(\xi, a) = M[L(\vartheta, a)] = \sum_{\vartheta \in \Theta} L(\vartheta, a)\, \xi(\vartheta). \qquad (9\text{-}2)$$

La mejor operación para el estadista será la llamada operación *de Bayes* a^*, con la cual las pérdidas serán mínimas e iguales a:

$$R^*(\xi) = L(\xi, a^*) = \min_{a \in A} L(\xi, a). \qquad (9\text{-}3)$$

El estadista no debe limitarse forzosamente a utilizar sólo una estrategia pura. El puede utilizar una mezcla de estrategias puras en correspondencia con cierta ley probabilística de distribución. En este caso vamos a referirnos a la *estrategia mixta del estadista.* Para emplear la estrategia mixta el estadista debe presentar la distribución de las probabilidades $\eta(a) = (\eta^{(1)}, \ldots, \eta^{(m)})$, que determina las probabilidades con las cuales va a utilizar sus estrategias puras $a_1, \ldots, a_m$. En el caso general dispone de cierta co-

lección de estrategias mixtas $H = \{\eta_1(a), \ldots, \eta_\nu(a)\}$ denominada *espacio de estrategias mixtas del estadista.*

Si el estadista emplea la estrategia mixta $\eta(a)$ y la naturaleza emplea la estrategia mixta $\xi(\vartheta)$, entonces las pérdidas medias del primero son

$$L(\xi, \eta) = M_{\vartheta, a}[L(\vartheta, a)] = \sum_{\vartheta, a} L(\vartheta, a)\,\xi(\vartheta)\,\eta(a). \qquad (9\text{-}4)$$

En este caso la tarea del estadista consiste en elegir una estrategia $\eta^*(a) \in H$ tal que con ella sus pérdidas medias $L(\xi, \eta^*)$ sean mínimas, es decir,

$$L(\xi, \eta^*) = \min_{\eta \in H} L(\xi, \eta). \qquad (9\text{-}5)$$

En los casos examinados se resolvió un problema estadístico relativamente sencillo, la determinación de la mejor estrategia del estadista basándose solamente en la información apriorística sobre el estado de la naturaleza. Aquí, el estadista no hace intentos de precisar sus conocimientos sobre el estado de la naturaleza efectuando un experimento. Por eso, el tipo dado de juego estadístico puede ser llamado *juego estadístico sin experimento.*

d) Ejemplos de juegos estadísticos

Para una emejor comprensión de la estructura de los juegos estadísticos y de los métodos de su resolución, pondremos varios ejemplos con los cuales más adelante vamos a ilustrar los postulados fundamentales de la teoría de las resoluciones estadísticas.

Ejemplo 9-1. Problema de sustitución del equipo. El equipo complejo y caro, instalado en una empresa, después de k años de trabajo puede encontrarse en uno de los tres estados:

ϑ_1, el equipo está en perfecta capacidad de funcionamiento y sólo requiere una pequeña reparación corriente;

ϑ_2, algunas piezas se han desgastado considerablemente y requieren una reparación seria o su cambio;

Tabla 9-1

Probabilidades apriorísticas de los estados de la naturaleza y pérdidas en el problema de sustitución del equipo

ϑ	$\xi(\vartheta)$	A		
		a_1	a_2	a_3
ϑ_1	0,2	1	3	5
ϑ_2	0,5	5	2	4
ϑ_3	0,3	7	6	3

ϑ_3, las piezas principales se han desgastado tanto que es imposible la explotación ulterior del equipo.

La experiencia del pasado, al explotar un equipo análogo, muestra que en el 20% de los casos éste puede encontrarse en el estado ϑ_1, en el 50% de las ocasiones, en el estado ϑ_2 y en el 30% de los casos, en el estado ϑ_3.

Para la empresa son posibles tres diferentes modos de proceder:

a_1, dejar funcionando el equipo un año más, realizando una pequeña reparación por sus propios medios;

a_2, realizar la reparación del equipo llamando a una brigada especializada de reparadores;

a_3, cambiar el equipo por un nuevo.

En la tabla 9-1 se dan las pérdidas que sufre la empresa con los distintos modos de proceder. Forman parte de la magnitud de las pérdidas el costo de la reparación o de la sustitución del equipo, así como las pérdidas motivadas por el empeoramiento de la calidad de la producción y por los paros provocados por el equipo defectuoso. En esta misma tabla se dan las probabilidades apriorísticas de los diversos estados de la naturaleza, es decir, la estrategia mixta de la misma $\xi(\vartheta)$.

Para la estrategia mixta presentada $\xi(\vartheta)$ las pérdidas medias con los distintos modos de proceder son:

$$L(\xi, a_1) = \sum_{\vartheta} L(\vartheta, a_1)\,\xi(\vartheta) = 1 \cdot 0{,}2 + 5 \cdot 0{,}5 + 7 \cdot 0{,}3 = 4{,}8;$$

$$L(\xi, a_2) = 3{,}4; \qquad L(\xi, a_3) = 3{,}9.$$

Ejemplo 9-2. Problema de la línea tecnológica. A la línea tecnológica puede llegar materia prima con pequeña cantidad de impurezas (ϑ_1) y con gran cantidad de impurezas (ϑ_2). Se sabe que en promedio llega un 60% de materia prima de primera clase y un 40% de segunda clase. Para utilizar materias primas de distintas clases se han previsto tres regimenes de funcionamiento de la línea tecnológica: a_1, a_2 y a_3. En la tabla 9-2 se dan las probabilidades apriorísticas de los estados de la naturaleza y las pérdidas, que reflejan la calidad de la producción elaborada y los gastos de materias primas según la calidad de éstas y el régimen de funcionamiento de la línea tecnológica.

Tabla 9-2

Probabilidades apriorísticas de los estados de la naturaleza y pérdidas en el problema de la línea tecnológica

ϑ	$\xi(\vartheta)$	A		
		a_1	a_2	a_3
ϑ_1	0,6	0	1	3
ϑ_2	0,4	5	3	2

Las pérdidas medias correspondientes a las probabilidades apriorísticas presentadas con diferentes regimenes de funcionamiento son:

$$L(\xi, a_1) = 2{,}0; \qquad L(\xi, a_2) = 1{,}8; \qquad L(\xi, a_3) = 2{,}6.$$

El establecimiento del régimen de funcionamiento a_2 será la operación de Bayes.

9-2. JUEGOS ESTADISTICOS SIN EXPERIMENTO

a) Representación del juego estadístico sin experimento en forma de juego S

El juego estadístico puede representarse en forma de juego S equivalente al igual que se hizo con los juegos estratégicos. Para esto vinculamos con cada una de las estrategias puras $a_j \times \times (j = 1, \ldots, l)$ el punto $C_j = (q_{1j}, \ldots, q_{mj})$ en el espacio de m dimensiones cuyas coordenadas serán las pérdidas del estadista $L(\vartheta_i, a_j) = q_{ij}$ con diferentes estados de la naturaleza $\vartheta_i (i = 1, \ldots, m)$. Así pues, en el problema de la linea tecnológica, a las estrategias puras del estadista a_1, a_2 y a_3 van a corresponder los puntos del plano $C_1 = (0, 5)$; $C_2 = (1, 3)$ y $C_3 = (3, 2)$, como se muestra en la figura 9-1. La envoltura convexa S^* de los puntos $\{C_1, C_2, C_3\}$ proporciona el dominio de todas las estrategias posibles tanto puras como mixtas del estadista.

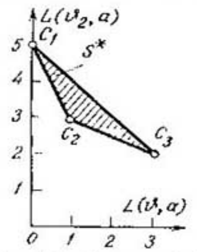


Fig. 9-1. Representación del problema de la línea technológica en forma de juego S

En lo adelante examinaremos varios principios por los cuales puede guiarse el estadista al alegir su estrategia. Al hacerlo, es necesario destacar que entre los estadistas no existe una opinión única en cuanto a cuál de los principios es el mejor en los juegos estadísticos. En otras palabras, no existe una regla universal que permita elegir una determinada forma de acción independientemente de la situación creada. Sin embargo, aunque pueden haber desacuerdos respecto a lo que debe hacerse en una situación dada llegamos a un acuerdo total respecto a lo que no ha de hacerse. Esto puede realizarse introduciendo el concepto de estrategias admisibles, análogo al de estrategias predominantes en los juegos estratégicos.

b) Estrategias admisibles en los juegos estadísticos

Supongamos, que examinamos la estrategia mixta del estadista $\eta(a)$. Pueden presentarse dos casos:

1) No puede hallarse ninguna estrategia mejor que $\eta(a)$. Esto significa que no existe una estrategia $\eta'(a)$ tal que

$$L(\vartheta, \eta') \leqslant L(\vartheta, \eta) \tag{9-6}$$

para todas las $\vartheta \in \Theta$, aunque para ciertas ϑ la relación (9-6) será válida. En este caso la estrategia $\eta(a)$ puede llamarse *admisible*. Pero ella no es necesariamente preferible ya que pueden haber otras estrategias que también merecen atención.

2) Existe la estrategia $\eta'(a)$, mejor que $\eta(a)$. Esto significa que para la estrategia $\eta'(a)$ la relación (9-6) va a cumplirse con todas las $\vartheta \in \Theta$. En este caso es necesario dejar de considerar la estrategia $\eta(a)$ en favor de la estrategia $\eta'(a)$, o sea, considerarla *inadmisible.*

Es cómodo examinar las estrategias admisibles en términos del juego S. Por cuanto en el juego S la estrategia del estadista se determina por el punto S de la envoltura convexa S^* y las pérdidas con las distintas $\vartheta \in \Theta$ se determinan por las coordenadas de este punto, la estrategia determinada por el punto S será admisible si no existe otro punto $S' \in S^*$, todas las coordenadas del cual serán menores que las correspondientes al punto S.

Examinemos el método de búsqueda de las estrategias admisibles para el caso en que el espacio de los estados de la naturaleza consta de los dos elementos ϑ_1 y ϑ_2. En la figura 9-2 se muestra el dominio convexo S^* correspondiente a este caso.

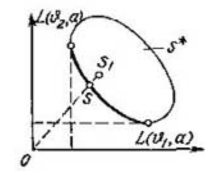


Fig. 9-2. Estrategias admisibles en el juego S

Analicemos la estrategia determinada por el punto S_1 que constituye un punto interior del dominio S^*. Dicha estrategia no es admisible, ya que todos los puntos que se encuentran en el segmento OS_1 dentro del S^* determinan estrategias mejores que S_1. La mejor de ellas es la estrategia S que pertenece al límite inferior izquierdo del dominio S^*. Por eso, todos los puntos interiores pueden eliminarse en favor de los puntos pertenecientes al límite inferior izquierdo del dominio S^* marcado en la figura 9-2 con una línea gruesa.

No obstante, el desplazamiento del punto a lo largo de este límite no da ventajas algunas ya que en este caso disminuyen las pérdidas correspondientes a un estado de la naturaleza pero aumentan las pérdidas correspondientes al otro estado de la naturaleza. Por eso los puntos pertenecientes al límite inferior izquierdo del dominio S^* determinan las estrategias admisibles del estadista.

Ejemplo 9-3. En el problema de la línea tecnológica (véase la fig. 9-1) el límite inferior izquierdo del dominio S^* consta de los segmentos C_1C_2 y C_2C_3, cada uno de los cuales se determina por la combinación de las estrategias puras a_1, a_2 y a_2, a_3. Sean $0 \leqslant w \leqslant 1$. Entonces, la ecuación para el segmento C_1C_2 se escribe en forma vectorial del modo siguiente:

$$S = wC_1 + (1 - w)\,C_2. \tag{9-7}$$

Esta ecuación define la estrategia mixta $\eta(a) = (w, 1 - w, 0)$. Al proyectar en los ejes de coordenadas la ecuación del segmento C_1C_2, obtenemos:

$$\left.\begin{aligned} L(\vartheta_1\eta) &= 0 \cdot w + 1\,(1 - w) = 1 - w; \\ L(\vartheta_2\eta) &= 5w + 3\,(1 - w) = 3 + 2w. \end{aligned}\right\} \tag{9-8}$$

Análogamente, la ecuación del segmento C_2C_3 que determina la estrategia mixta $\eta(a) = (0, w, 1-w)$ se reduce a la forma

$$L(\vartheta_1, \eta) = 3 - 2w, \quad L(\vartheta_2, \eta) = 2 + w. \qquad (9\text{-}9)$$

c) Principio de elección de las estrategias en los juegos estadísticos

Se denomina principio de elección la regla que permite determinar la mejor estrategia mixta del estadista. En diversos casos el estadista puede utilizar diferentes principios de su estrategia.

El *principio de minimax* puede ser uno de los posibles principios de elección de la estrategia. Este principio se emplea con éxito en los juegos estratégicos cuando se juega contra un adversario inteligente que desea hacernos el daño máximo. No obstante, en diferentes casos también es útil emplear este principio en los juegos estadísticos.

Según el principio de minimax el estadista elige una estrategia mixta $\eta(a)$ con la cual las pérdidas medias $L(\vartheta, \eta)$ van a ser mínimas siendo el estado de la naturaleza ϑ el peor para éste. El caso peor será aquel $\vartheta \in \Theta$ en que la magnitud $L(\vartheta, \eta)$ adquiere el valor máximo. El estadista debe minimizar esta magnitud, o sea, elegir la estrategia $\eta^*(a)$ que garantiza el cumplimiento de la condición

$$L(\vartheta, \eta^*) = \min_{\eta} \max_{\vartheta} L(\vartheta, \eta). \qquad (9\text{-}10)$$

Ejemplo 9-4. Hallemos la estrategia de minimax del estadista en el problema de la línea tecnológica. Como la solución debe hallarse en la clase de estrategias admisibles, es suficiente limitarse al análisis de las estrategias definidas por las relaciones (9-8) y (9-9) y que corresponden a los segmentos C_1C_2

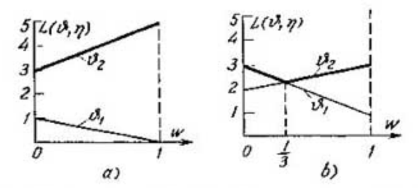


Fig. 9-3. Determinación de la estrategia óptima por el principio de minimax (*a*, corresponde al segmento C_1C_2 y *b*, al segmento C_2C_3 en la fig. 9-1)

y C_2C_3 de la envoltura convexa S^* dada en la figura 9-1. Conforme a estas relaciones y segmentos, en la figura 9-3, *a* y *b* se han construido las gráficas $L(\vartheta, \eta)$ en la función de w para $\vartheta = \vartheta_1$ y $\vartheta = \vartheta_2$. Los valores $\max_{\vartheta} L(\vartheta, \eta)$ están marcados con líneas gruesas. Como se ve en las figuras, el mínimo de esta magnitud se alcanza igual a 3 en el segmento C_1C_2, y en el segmento C_2C_3 se determina por la condición de que se intersecan dos rectas

$$3 - 2w = 2 + w,$$

es decir, se alcanza con $w = 1/3$ y es igual a $7/3 < 3$. De este modo, el principio de minimax da un punto en el segmento C_2C_3 correspondiente a $w = 1/3$, o sea, determina la estrategia mixta óptima $\eta^* = (0, 1/3, 2/3)$ con la cual las pérdidas del estadista no serán mayores que la magnitud $7/3$ con cualquiera estrategia de la naturaleza.

A veces es racional elegir la estrategia no partiendo de las pérdidas totales $L(\vartheta, a)$ sino de las llamadas *pérdidas complementarias* $L'(\vartheta_1 a)$ que se definen por las relaciones

$$L'(\vartheta, a) = L(\vartheta, a) - \min_a L(\vartheta, a). \quad (9\text{-}11)$$

Para cada estado de la naturaleza la magnitud $\min\limits_a L(\vartheta, a)$ determina las pérdidas mínimas que el estadista sufre obligatoriametne, incluso con su mejor operación, es decir, las pérdidas necesarias. Puede considerarse que las pérdidas necesarias se compensan de una u otra forma (por ejemplo, fijando los precios adecuados de la producción) y pueden dejar de tomarse en cuenta al elegir la estrategia. En este caso la elección de la estrategia puede realizarse según el principio *de minimax de las pérdidas complementarias*.

Ejemplo 9-5. Las pérdidas complementarias en el problema de la línea tecnológica, halladas mediante la relación (9-11) se han resumido en la tabla 9-3. Aplicando a los datos de esta tabla el principio de minimax, obtenemos en calidad de estrategia óptima la estrategia pura a_2 con las pérdidas 1.

Tabla 9-3

Pérdidas complementarias en el problema de la línea tecnológica

ϑ	A		
	a_1	a_2	a_3
ϑ_1	0	1	3
ϑ_2	3	1	0

Los principios de minimax que parten de la suposición de que la naturaleza obra del peor modo para el estadista, están justificados en los juegos estratégicos, pero en los juegos estadísticos expresan realmente el punto de vista de una persona muy cuidadosa que trata de obtener aunque sea algo de lo alcanzable y no persigue el máximo para no sufrir grandes pérdidas aleatorias. Es necesario considerar también una desventaja de los principios de minimax debido a que éstos no tienen en cuenta la información apriorística sobre el estado de la naturaleza y con ello limitan la ganancia que puede proporcionar dicha información.

Por eso, los principos de minimax pueden recomendarse en los casos en que falta información apriorística sobre los estados de la naturaleza o existe fundamento para dudar de que la información sea cierta.

Otro principio de elección de estrategia que toma en consideración la distribución apriorística de las probabilidades $\xi(\vartheta)$ es el *de Bayes*. Según este principio la estimación de la estrategia mixta del estadista $\eta(a)$ se realiza promediando las pérdidas $L(\vartheta, \eta)$

por todos los estados posibles de la naturaleza $\vartheta \in \Theta$ teniendo en cuenta la distribución apriorística de las probabilidades $\xi(\vartheta)$, es decir, por la magnitud

$$L(\xi, \eta) = \sum_{\vartheta} L(\vartheta, \eta)\, \xi(\vartheta). \tag{9-12}$$

Por cierto, la mejor estrategia $\eta(a)$ será aquella que proporciona el mínimo de la magnitud $L(\xi, \eta)$. Esta estrategia óptima se denomina *estrategia de Bayes.*

Naturalmente, el principio de Bayes puede aplicarse tanto a las pérdidas totales como a las complementarias. Sin embargo, en los ejemplos ulteriores nos limitaremos solamente a aplicar el principio de Bayes a las pérdidas totales.

Ejemplo 9-6. En el problema de la línea tecnológica con $\xi(\vartheta_1) = 0{,}6$ y $\xi(\vartheta_2) = 0{,}4$, para las estrategias admisibles determinadas por el segmento C_1C_2, tenemos:

$$L(\xi, \eta) = (1 - w)\, 0{,}6 + (3 + 2w)\, 0{,}4 = 1{,}8 + 2{,}0\, w;$$

de manera que $\min\limits_{w} L(\xi, \eta) = 1{,}8$ con $w = 0$, lo que corresponde a la estrategia mixta $\eta = (0, 1, 0)$; para el segmento C_2C_3

$$L(\xi, \eta) = 2{,}6 - 0{,}8w,$$

de suerte que $\min\limits_{w} L(\xi, \eta) = 1{,}8$ para $w = 1$, lo que corresponde a la misma estrategia mixta $\eta = (0, 1, 0)$. De este modo, la estrategia pura a_2 es una estrategia de Bayes.

Es necesario notar que el resultado obtenido no es aleatorio. Más adelante mostraremos que el principio de Bayes siempre brinda en calidad de estrategia óptima una de las estrategias puras del estadista. Entonces la operación de Bayes a^* se determinará por la condición (9-3).

d) Interpretación geométrica de las estrategias de Bayes

Examinemos el juego estadístico S para los dos estados de la naturaleza ϑ_1 y ϑ_2, determinado por la zona convexa S^* en el plano (x, y) mostrado en la figura 9-4, donde

$$x = L(\vartheta_1, \eta); \quad y = L(\vartheta_2, \eta). \tag{9-13}$$

Como sabemos, las estrategias admisibles se hallan en el límite inferior izquierdo de la zona S^*. Examinemos una de las estrategias S_0. Construyamos el conjunto auxiliar R que contiene todos los puntos que se encuentran más abajo y a la izquierda del punto S_0. Evidentemente que S^* y R son conjuntos convexos y ningún punto de S^*, incluso el punto S_0, pertenece a R.

Tracemos la recta que divide los conjuntos S^* y R. Esta recta debe pasar a través de S_0, es decir, debe ser la recta de referencia al conjunto S^* en el punto S_0. Pero ella debe ser también recta de referencia para el conjunto R. Luego, dicha recta deberá tener pen-

diente negativa, ser vertical u horizontal, y su ecuación puede escribirse en la forma

$$y = -kx + c, \quad k \geqslant 0. \tag{9-14}$$

Al dividir ambos miembros de ésta ecuación por $k+1$ la reducimos a la forma

$$ax + by = c', \tag{9-15}$$

donde

$$a = \frac{k}{k+1}; \quad b = \frac{1}{k+1}; \quad c' = \frac{c}{k+1}. \tag{9-16}$$

Notando que $a \geqslant 0$, $b \geqslant 0$, $a + b = 1$ podemos hacer $a = w$, $b = 1 - w$ y considerar w y $1 - w$ como probabilidades apriorísticas de los estados de la naturaleza ϑ_1 y ϑ_2:

$$w = \xi(\vartheta_1), \quad 1 - w = \xi(\vartheta_2). \tag{9-17}$$

En este caso la ecuación de la recta de referencia puede escribirse en la forma

$$wx + (1 - w)\,y = c' \tag{9-18}$$

o bien

$$L(\vartheta_1, \eta)\,\xi(\vartheta_1) + L(\vartheta_2, \eta)\,\xi(\vartheta_2) = c'. \tag{9-19}$$

Como vemos, la magnitud c' determina las pérdidas medias $L(\xi, \eta)$ con las probabilidades apriorísticas de los estados de la naturaleza $\xi(\vartheta_1) = w$ y $\xi(\vartheta_2) = 1 - w$. Es fácil notar asimismo, que la magnitud c' para el punto S_0 es el mínimo de todos los valores posibles de $L(\xi, \eta)$, ya que el aumento de c' hasta $c'' > c'$ significa que la recta pasará más arriba y va a pasar por los puntos del conjunto S^* que no son admisibles y la disminución de c' no es posible, ya que la recta descenderá y no va a tener puntos comunes con S^*. De este modo la magnitud c' determina la estrategia η que da un mínimo de pérdidas medias $L(\xi, \eta)$ con la distribución dada $\xi(\vartheta) = (w, 1 - w)$, es decir, determina la estrategia de Bayes para la distribución apriorística de probabilidades dada.

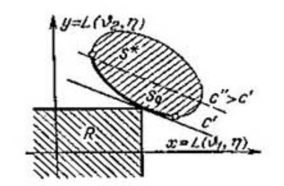


Fig. 9-4. Construcción de la línea de apoyo al conjunto convexo

Por cuanto S_0 constituye un punto arbitrario en el límite de las estrategias admisibles, para cualquier punto de este límite pueden encontrarse tales w y $1 - w$ para las cuales este punto va a determinar la estrategia de Bayes. De aquí se deduce que *cada estrategia admisible constituye la estrategia de Bayes para ciertas probabilidades apriorísticas* w *y* $1 - w$.

Supongamos ahora que nos han presentado las probabilidades apriorísticas de los estados de la naturaleza w y $1 - w$ y se requiere hallar el punto en la envoltura convexa S^* que determina

la estrategia de Bayes para este caso. Tracemos en el plano (x, y) la recta

$$wx + (1 - w)\,y = c. \tag{9-20}$$

Con una c arbitraria esta recta será paralela a la recta de referencia en el punto correspondiente a la estrategia de Bayes. Para facilitar el trazado, supongamos que $c = w(1 - w)$, y escribamos la ecuación (9-20) en la forma

$$\frac{x}{1 - w} + \frac{y}{w} = 1, \tag{9-21}$$

que representa la ecuación de una recta en segmentos. En la figura 9-5 está trazada la recta correspondiente a la ecuación (9-21). Ahora ya es fácil trazar la recta de referencia correspondiente a los valores dados de w y $1 - w$ que determina en la envoltura convexa S^* el punto correspondiente al a estrategia de Bayes del estadista.

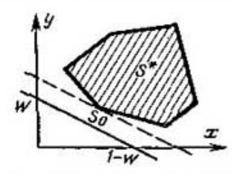


Fig. 9-5. Determinación geométrica de la estrategia de Bayes

Por cuanto el límite exterior de la envoltura convexa es un polígono con los véritces correspondientes a las estrategias puras del estadista, la recta de referencia pasa obligatoriamente por uno siquiera de los vértices. Luego, *para las probabilidades aprioristicas dadas w y $1 - w$ siempre existe al menos una estrategia de Bayes que es pura.* Esta circunstancia simplifica muchísimo la resolución de los juegos estadísticos ya que al buscar la solución de Bayes, permite limitarse a examinar sólo un número finito de estrategias puras admisibles en lugar de un conjunto infinito de estrategias mixtas. Esto satisface también a muchos estadistas de ánimo escéptico que están en contra del empleo de las estrategias mixtas fundándose en que con ellos se requiere utilizar un mecanismo de elección aleatoria que no guarda relación con la esencia del problema.

Por supuesto que todas las deducciones hechas siguen siendo válidas cuando existen más de dos estados de la naturaleza. Pero entonces ya es imposible el análisis en el plano lo que dificulta la ilustración geométrica de estos casos.

9-3. JUEGOS ESTADISTICOS CON REALIZACION DE UN EXPERIMENTO UNICO

a) Enunciado del problema

Como se ha señalado ya la peculiaridad de los juegos estadísticos es que el estadista puede profundizar y precisar sus conocimientos respecto al estado de la naturaleza efectuando un experi-

mento. En principio, si el estadista tuviera la posibilidad ilimitada de experimentar, obtendría información completa sobre el estado de la naturaleza y obraría en condiciones de plena determinación. No obstante, el montar un experimento siempre acarrea gastos de recursos y tiempo, cuyas pérdidas pueden resultar más cuantiosas que el beneficio obtenido del experimento.

La posibilidad de realizar el experimento amplía extraordinariamente la clase de las estrategias del estadista. Ante todo, el estadista debe tomar la decisión de efectuar o no efectuar el experimento. Luego, tiene que decidir de qué naturaleza debe ser este experimento, cuántas pruebas diferentes será necesario realizar antes de considerar terminado el experimento y qué medidas hay que tomar al obtener unos u otros resultados del experimento.

Trataremos de responder a algunas de estas interrogaciones en los párrafos siguientes. En el párrafo dado centraremos nuestra atención en los juegos estadísticos con experimento único.

Por ahora no vamos a incluir en la clase de estrategias del estadista la toma de decisión sobre la realización del experimento considerando que ya se ha tomado la decisión de efectuar el experimento único. Por cierto, vamos a entender por experimento único aquel cuya extensión y orden de realización se han determinado previamente. Así pues, en caso de ser necesario comprobar si es simétrica una moneda dada, puede efectuarse el experimento único consistente en lanzar n veces la moneda. El resultado del experimento con $n = 5$ puede ser la secuencia *JPIJP*. Existen solamente 2^n de estas secuencias, es decir, en el caso dado el espacio de resultados del experimento consta de 2^n elementos. Análogamente, al comprobar un arma se entiende por experimento único aquel durante el cual se hacen n disparos. Para estimar cómo influye sobre un animal cierto tipo especial de alimento, puede realizarse el experimento único consistente en medir durante varios meses el aumento de peso diario en n animales, etc. Aunque en estos ejemplos el experimento consta de diferentes subpruebas, lo denominamos único ya que el número de subpruebas y el carácter de cada una están determinados de antemano.

b) Espacio muestral

Designemos por Z el espacio de resultados del experimento. Los elementos de dicho espacio, o sea, los distintos resultados del experimento serán $z_1, \ldots, z_v$.

Los distintos resultados del experimento $z \in Z$ se hallan relacionados con los estados de la naturaleza $\vartheta \in \Theta$. Esta relación consiste en que para cada estado de la naturaleza $\vartheta \in \Theta$ hay una probabilidad determinada $p_\vartheta(z)$ de que el resultado del experimento será la $z \in Z$ dada. Las magnitudes $p_\vartheta(z)$, que a veces se designan como $p(z|\vartheta)$, representan la distribución convencional

de probabilidades en el espacio Z con la ϑ dada y obedecen a las relaciones

$$p_\vartheta(z) \geqslant 0; \quad \sum_z p_\vartheta(z) = 1. \tag{9-22}$$

El total de los tres elementos: espacio de resultados del experimento Z, espacio de estados de la naturaleza Θ y distribución de probabilidades $p_\vartheta(z)$ en el espacio Z con $\vartheta \in \Theta$ dada se denomina *espacio muestral* y se designa por

$$\mathfrak{Z} = (z, \Theta, p). \tag{9-23}$$

Es cómodo presentar el espacio muestral en forma de tabla que contiene la distribución $p_\vartheta(z)$ en el producto directo de los conjuntos $\Theta \times z$.

Ejemplo 9-7. En el problema de sustitución del equipo el experimento puede consistir en realizar las pruebas del equipo por los medios de la empresa. En este caso la insuficiente calificación del personal y falta de los aparatos de medición requeridos provocan que los resultados de las pruebas sólo reflejen aproximadamente el estado del equipo. Supongamos que el experimento puede tener cuatro resultados:

z_1, el equipo está en buenas condiciones;
z_2, se requiere una reparación corriente;
z_3, se requiere cambiar las piezas desgastadas;
z_4, el equipo no sirve para su explotación ulterior.

Las probabilidades de cada uno de estos resultados, en distintos estados de la naturaleza, se ofrecen en la tabla 9-4 que representa el espacio muestral para el problema considerado.

Ejemplo 9-8. En el problema de la línea tecnológica el experimento puede consistir en el análisis preliminar aproximado de las impurezas contenidas (no es posible realizar un análisis de laboratorio preciso, ya que requiere grandes pérdidas de tiempo lo que significa horas muertas para el equipo). Los resultados del experimento son: z_1, no se han descubierto impurezas; z_2, hay pocas impurezas; z_3, hay muchas impurezas. Entonces, el espacio muestral puede tener el aspecto de la tabla 9-5.

Tabla 9-4

Espacio muestral en el problema de sustitución del equipo

ϑ	z			
	z_1	z_2	z_3	z_4
ϑ_1	$1/2$	$1/2$	0	0
ϑ_2	0	$1/2$	$1/2$	0
ϑ_3	0	0	$1/3$	$2/3$

Tabla 9-5

Espacio muestral en el problema de la línea tecnológica

ϑ	z		
	z_1	z_2	z_3
ϑ_1	0,60	0,25	0,15
ϑ_2	0,20	0,30	0,50

c) Función de decisión

En el problema sin experimento el estadista debía tomar una decisión del espacio de decisiones A, basándose en la información apriorística $\xi(\vartheta)$ sobre los estados de la naturaleza. En el pro-

blema con experimento el mismo toma la decisión según el resultado del experimento $z \in Z$. Para formalizar este problema, él puede analizar por anticipado todos los resultados posibles del experimento y formular la regla d que determina cuál decisión $a \in A$ debe tomarse con cada resultado posible del experimento $z \in Z$. Esta regla representará el reflejo del espacio de resultados del experimento Z en el espacio de decisiones A

$$d : Z \rightarrow A, \tag{9-24}$$

lo que puede escribirse también en la forma

$$d(z) = a.$$

Se denomina *función de decisión* la regla $d(z)$ que determina la decisión $a \in A$ que debe tomar el estadista con cualquier resultado del experimento $z \in Z$.

Aclaremos el concepto de función de decisión con el ejemplo siguiente. Supongamos que el espacio de decisiones consta de tres elementos $A = \{a_1, a_2, a_3\}$ y el espacio de resultados del experimento, de cinco $Z = \{z_1, z_2, z_3, z_4, z_5\}$. La función de decisión $d(z_i) = a_i$ puede presentarse en forma de conjunto de pares de subíndices (i, j), que determinan la decisión a_j con el resultado del experimento z_i. La función de decisión será, por ejemplo, el conjunto $\{(1, 1), (2, 1), (3, 2), (4, 2), (5, 3)\}$; que significa que con los resultados z_1 y z_2 se toma la decisión a_1, con los resultados z_3 y z_4 se asume la decisión a_2 y con el resultado z_5 se toma la decisión a_3. Claro que la función de decisión dada no es la única posible. Sería posible asimismo examinar las funciones de decisión del tipo $\{(1, 1), (2, 2), (3, 2), (4, 3), (5, 3)\}$ ó $\{(1, 3), (2, 1), (3, 2), (4, 3), (5, 2)\}$, etc. De acuerdo con esto conviene introducir el espacio D que contiene la enumeración completa de las funciones de decisión posibles y denominarlo *espacio de funciones de decisión*.

La función de decisión examinada más arriba fracciona el conjunto Z en subconjuntos no intersecados S_{a_i}, a saber:

$$S_{a_1} = \{z_1, z_2\}, \quad S_{a_2} = \{z_3, z_4\}, \quad S_{a_3} = \{z_5\}. \tag{9-25}$$

En términos de la teoría de conjuntos puede decirse que es posible considerar cualquier función de decisión $d \in D$ como fraccionamiento del espacio de resultados del experimento Z en subconjuntos no intersecados S_a tales que

$$S_a = \{z : d(z) = a\}, \quad a \in A. \tag{9-26}$$

En particular, si el espacio de decisiones A consta sólo de dos elementos a_1 y a_2 (problema de dos alternativas), la función de decisión $d(z)$ fracciona el espacio Z en el S, llamado dominio crítico y en su complemento $C(S) = \bar{S}$ que se determina por las condiciones

$$d(z) = \begin{cases} a_1, & \text{si } z \in S; \\ a_2, & \text{si } z \in C(S). \end{cases} \tag{9-27}$$

El concepto de función de decisión posibilita formular más precisamente la tarea del estadista. Esta tarea consiste en escoger del espacio de funciones de decisión D aquella función $d(z)$ que permite tomar las decisiones más ventajosas. Pero para esto es necesario saber estimar las diferentes funciones de decisión lo que puede hacerse con ayuda de la función de riesgo.

d) Función de riesgo

Si el estadista realizó la elección de cierta función de decisión $d(z)$, con esto él determinó para cada resultado del experimento $z \in Z$ la decisión $a = d(z)$ a la que con $\vartheta \in \Theta$ van a conrresponder las pérdidas

$$L(\vartheta, a) = L[\vartheta, d(z)] = L_z(\vartheta, d). \qquad (9\text{-}28)$$

No obstante, con una ϑ dada el resultado del experimento z será una magnitud aleatoria definida por la distribución de probabilidades $p_\theta(z)$ en el espacio Z. Luego, para la z dada van a ocurrir con la misma probabilidad $p_\theta(z)$ las pérdidas $L_z(\vartheta, d)$ que de este modo serán también una magnitud aleatoria.

Por cuanto, para estimar la función de decisión $d(z)$ con el estado dado de la naturaleza ϑ hay que tomar en cuenta todos los resultados posibles del experimento, es necesario referirnos a las pérdidas medias determinadas en todo el espacio de resultados del experimento Z. Estas pérdidas medias se denominan *función de riesgo*, se designan por $\rho(\vartheta, d)$ y se determinan por la relación

$$\rho(\vartheta, d) = M_z[L_z(\vartheta, d)] = \sum_z L_z(\vartheta, d)\, p_\theta(z). \qquad (9\text{-}29)$$

A cada estado de la naturaleza $\vartheta \in \Theta$ y cada función de decisión $d \in D$ les corresponderá su valor de pérdidas medias, es decir, la función de riesgo $\rho(\vartheta, d)$ que, de este modo, se determinará en el producto directo de los conjuntos $\Theta \times D$ en forma completamente análoga a como la función de pérdidas $L(\vartheta, a)$ en el juego sin experimento se determinó en el producto directo de los conjuntos $\Theta \times A$. De esto se deduce que el espacio de las funciones de decisión D y la función de riesgo $\rho(\vartheta, d)$ en el juego con experimento único juegan el mismo papel que el espacio de decisiones A y la función de pérdidas en el juego sin experimento. De aquí se infieren métodos análogos para resolver estos dos problemas.

En los juegos con experimento único el estadista puede usar también estrategias mixtas. Para esto debe disponer de un mecanismo de elección aleatoria que determine la distribución de probabilidades $\eta(d)$ en el espacio D. La función de riesgo, al utilizar la estrategia mixta $\eta(d)$ designada en este caso por $\rho(\vartheta, \eta)$, se obtiene promediando $\rho(\vartheta, d)$ por todas las estrategias puras que forman parte de la estrategia mixta dada. De este modo,

$$\rho(\vartheta, \eta) = M_d[\rho(\vartheta, d)] = \sum_d \rho(\vartheta, d)\, \eta(d) \qquad (9\text{-}30)$$

Tabla 9-6

Pérdidas y probabilidades de resultados del experimento en el problema de la línea tecnológica

ϑ	a			$p_\vartheta(z)$		
	a_1	a_2	a_3	z_1	z_2	z_3
ϑ_1	0	1	3	0,60	0,25	0,15
ϑ_2	5	3	2	0,20	0,30	0,50

o tamando en consideración (9-29)

$$\rho(\vartheta, \eta) = \sum_{z,\ d} L_z(\vartheta, d)\, p_\vartheta(z)\, \eta(d). \tag{9-31}$$

Naturalmente, al buscar la mejor estrategia en el juego con experimento único el estadista debe partir solamente de las estrategias admisibles que se determinan exactamente al igual que en el juego sin experimento.

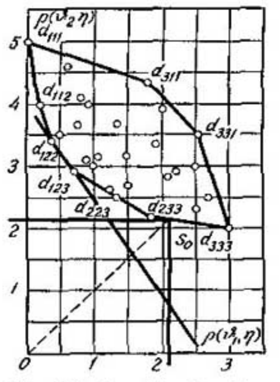


Fig. 9-6. Función de riesgo en el problema de la línea tecnológica

Ejemplo 9-9. Determinemos la función de riesgo en el problema de la línea tecnológica. Para facilitar los cálculos, reduciremos a una tabla (tabla 9-6) las pérdidas del estadista $L(\vartheta, a)$ y las probabilidades de los resultados del experimento $p_\vartheta(z)$ dadas en las tablas 9-2 y 9-5.

Por cuanto el espacio de resultados del experimento consta de tres elementos, la función de decisión va a tener la forma $d(z) = (a_i, a_j, a_k) = d_{ijk}$, donde a_i, a_j y a_k significan las decisiones que debe tomar el estadista con los resultados del experimento z_1, z_2 y z_3, respectivamente. Así, la función de decisión d_{122} significa que con los resultados del experimento z_1, z_2 y z_3 el estadista toma las decisiones a_1, a_2 y a_3.

Los valores de la función de riesgo, calculados por las fórmulas (9-29) para cada función de decisión, se exponen en la tabla 9-7 y la figura 9-6. Examinemos con el ejemplo del cálculo de $\rho(\vartheta, d_{122})$, los datos obtenidos de esta tabla. Según (9-29) tenemos:

$$\rho(\vartheta, d_{122}) = L_{z_1}(\vartheta, d_{122})\, p_\vartheta(z_1) + L_{z_2}(\vartheta, d_{122}) \times$$
$$\times p_\vartheta(z_2) + L_{z_3}(\vartheta, d_{122})\, p_\vartheta(z_3) = L(\vartheta, a_1) \times$$
$$\times p_\vartheta(z_1) + L(\vartheta, a_2)\, p_\vartheta(z_2) + L(\vartheta, a_2)\, p_\vartheta(z_3).$$

Suponiendo que $\vartheta = \vartheta_1$ y $\vartheta = \vartheta_2$, obtenemos:

$$\rho(\vartheta_1, d_{122}) = 0{,}4; \quad \rho(\vartheta_2, d_{122}) = 3{,}4.$$

Tabla 9-7

Valores de la función de riesgo $\rho(\vartheta, d)$ en el problema de la línea tecnológica

ϑ	d_{111}	d_{112}	d_{113}	d_{121}	d_{122}	d_{123}	d_{131}	d_{132}	d_{133}
ϑ_1	0,00	0,15	0,45	0,25	0,40	0,70	0,75	0,90	1,20
ϑ_2	5,00	4,00	3,50	4,40	3,40	2,90	4,10	3,10	2,60
Continuación de la tabla 9-7									
ϑ	d_{211}	d_{212}	d_{213}	d_{221}	d_{222}	d_{223}	d_{231}	d_{232}	d_{233}
ϑ_1	0,60	0,75	1,05	0,85	1,00	1,30	1,35	1,50	1,80
ϑ_2	4,60	3,60	3,10	4,00	3,00	2,50	3,70	2,70	2,20
Continuación de la tabla 9-7									
ϑ	d_{311}	d_{312}	d_{313}	d_{321}	d_{322}	d_{323}	d_{331}	d_{332}	d_{333}
ϑ_1	1,80	1,95	2,25	2,05	2,20	2,50	2,55	2,70	3,00
ϑ_2	4,40	3,40	2,90	3,80	2,80	2,30	3,50	2,50	2,00

e) Principios de elección de la estrategia en los juegos con experimento único

Por cuanto la introducción de la función de riesgo reduce el juego con experimento único a una forma análoga al juego sin experimento, todos los principios de elección de estrategia en el juego sin experimento se mantienen vigentes para el caso dado, con la diferencia de que en lugar de minimizar las pérdidas medias el estadista debe ahora minimizar el riesgo medio.

El principio de minimax consiste en elegir la estrategia $\eta(d)$ con la que el riesgo medio $\rho(\vartheta, \eta)$ sea el mínimo en el peor estado de la naturaleza para el estadista, es decir, la elección de la estrategia de minimax η^* se realiza por medio de la condición

$$\rho(\vartheta, \eta^*) = \min_{\eta} \max_{\vartheta} \rho(\vartheta, \eta). \qquad (9\text{-}32)$$

Al determinar el riesgo medio $\rho(\vartheta, \eta)$, si partimos de las pérdidas no completas sino de las complementarias $L'(\vartheta, a)$, la fórmula (9-32) definirá el principio de minimax de las pérdidas complementarias.

Para aplicar el principio de Bayes, introduzcamos el concepto de *riesgo esperado*, por el que se entiende el riesgo medio tomando en consideración todos los estados posibles de la naturaleza $\vartheta \in \Theta$ y la distribución apriorística de las probabilidades en el espacio Θ. Como que el estadista, al utilizar el principio de Bayes, puede limitarse a emplear solamente estrategias puras, el riesgo esperado es

$$\rho(\xi, d) = \sum_{\vartheta} \rho(\vartheta, d)\, \xi(\vartheta). \tag{9-33}$$

El principio de Bayes requiere el empleo de una función de decisión d^* con la que será mínimo el riesgo esperado, denominado en este caso *riesgo de Bayes*:

$$\rho^*(\xi) = \rho(\xi, d^*) = \min_{d} \rho(\xi, d). \tag{9-34}$$

Ejemplo 9-10. Determinemos las estrategias, de minimax y Bayes en el problema de la línea tecnológica con realización de experimento único. En la figura 9-6 este problema se ha presentado en forma del juego S. Gracias a la construcción geométrica dada resulta fácil establecer que la estrategia de minimax se determina por el punto S_0 con las coordenandas $({}^{30}/_{14}, {}^{30}/_{14})$ y corresponde al empleo de las estrategias puras d_{233} y d_{333} con las probabilidades ${}^{10}/_{14}$ y ${}^{4}/_{14}$. La estrategia d_{123} será la estrategia de Bayes.

9-4. UTILIZACION DE LAS PROBABILIDADES APOSTERIORISTICAS

a) Determinación del número de estrategias en los juegos con experimento

En los ejemplos de los párrafos anteriores se ve que en los juegos estadísticos la realización del experimento conduce al aumento considerable del número de estrategias puras del estadista. Así pues, en el problema de la línea tecnológica, al tomar en consideración los resultados del experimento, el número de estrategias puras creció de 3 a 27. Si no resultara posible representar este juego en el plano en forma de juego S surgirían serias dificultades en su análisis.

Designemos por N el número total de estrategias puras del estadista en el juego con experimento único. No es difícil calcular este número. Supongamos que en el juego sin experimento, el estadista puede tomar l decisiones posibles $a_1, \ldots, a_l$ y el número de resultados posibles del experimento es igual a ν.

Con $\nu = 1$ obtenemos un juego en el que se sabe por anticipado el resultado del experimento lo que es equivalente al juego sin experimento con $N = l$.

Con $\nu = 2$ la estrategia del estadista puede escribirse en la forma $d = (a_{i_1}, a_{i_2})$ en que las magnitudes a_{i_1} y a_{i_2} pueden ser cualesquier $a \in A$. Por consiguiente, puede haber l valores diferentes

de a_{l_1}, a cada uno de los cuales pueden corresponder l valores diferentes de a_{l}, de modo que $N = l^2$.

Con $\nu = 3$ la estrategia del estadista es $d = (a_{l_1}, a_{l_2}, a_{l_3})$. El total de soluciones a_{l_1} y a_{l_2} puede adoptar l^2 valores a cada uno de los cuales pueden corresponder l valores distintos de a_{l_3}, de suerte que $N = l^3$.

Al proseguir razonando análogamente llegamos a la conclusión de que con l y ν arbitrarias el número total de estrategias puras es de $N = l_{\nu}$. El número de estrategias puras del estadista con l y ν grandes puede resultar tan considerable que surgirán serias dificultades relacionadas con su análisis.

En tales casos podemos lograr una gran simplificación en la búsqueda de las estrategias de Bayes si en lugar de la distribución apriorística de probabilidades empleamos la aposteriorística calculada basándonos en los resultados del experimento efectuado. Con esto, el número de estrategias puras del estadista en el problema con experimento se mantendrá al igual que en el problema sin experimento.

b) Distribución aposteriorística de probabilidades. Fórmula de Bayes

La distribución apriorística de probabilidades $\xi(\vartheta)$, obtenida fundándose en los datos y la experiencia del pasado, brinda información útil sobre la frecuencia con que, en general, se presenta uno u otro estado de la naturaleza independientemente de las condiciones concretas en que el estadista tiene que tomar la decisión. El objetivo del experimento llevado a cabo por el estadista es recibir información adicional acerca del estado real de la naturaleza.

Supongamos que el espacio de resultados del experimento es el conjunto $Z = \{z_1, \ldots, z_\nu\}$. En este caso el resultado del experimento concreto será aleatorio y por eso tampoco permite determinar con precisión el verdadero estado de la naturaleza. No obstante, la indeterminación respecto al estado de la naturaleza disminuye notablemente si el experimento se ha realizado de forma correcta.

Este cambio de la indeterminación respecto al estado de la naturaleza consiste en que a resultas del experimento, en lugar de la distribución apriorística $\xi(\vartheta)$ se obtiene una nueva distribución de probabilidades $\xi_z(\vartheta)$ que se denomina *distribución aposteriorística de probabilidades* en el espacio Θ con el resultado concreto dado del experimento $z \in Z$. Ocupémonos de los métodos de cálculo de la distribución apriorística de probabilidades.

Ante todo, notemos que el resultado del experimento que tiene por objeto precisar el estado real de la naturaleza, dependerá del estado de la naturaleza ϑ. El estudio preliminar de las condiciones en las cuales se realiza el experimento permite indicar para cada estado de la naturaleza ϑ, la distribución de probabilidades en el

espacio de resultados del experimento Z que, de esta suerte, va a representar la distribución condicional de probabilidades $p(Z|\vartheta) = p_\vartheta(z)$.

Para describir de forma completa un experimento concreto hay que conocer el resultado del experimento $z \in Z$ y el estado de la naturaleza $\vartheta \in \Theta$ con el cual se realizó el mismo. Luego entonces, el resultado de un experimento concreto puede representarse en forma del par ordenado (z, ϑ), que constituye un elemento del producto directo de los conjuntos $Z \times \Theta$.

Designemos por $q(z, \vartheta)$ la distribución de probabilidades en el conjunto $Z \times \Theta$. En el caso general esta distribución de probabilidades va a variar al cambiar la distribución apriorística de probabilidades $\xi(\vartheta)$, es decir, será función de ξ. Para la distribución apriorística concreta $\xi(\vartheta)$ designaremos la distribución de probabilidades en el conjunto $Z \times \Theta$ por $q_\xi(z, \vartheta)$.

Nos es necesario relacionar entre sí las distribuciones $\xi(\vartheta)$, $p_\vartheta(z)$, $\xi_z(\vartheta)$ y $q_\xi(z, \vartheta)$. Ya hemos tropezado con un problema similar al examinar las magnitudes aleatorias bidimensionales en las que el resultado del experimento z era el par ordenado (x, y). Para estas magnitudes aleatorias se cumple la relación (5-37) que, al aplicarse a nuestro caso, puede escribirse en la forma

$$q_\xi(z, \vartheta) = p_\vartheta(z)\,\xi(\vartheta) = \xi_z(\vartheta)\,p(z). \tag{9-35}$$

De aquí hallamos que

$$\xi_z(\vartheta) = \frac{p_\vartheta(z)\,\xi(\vartheta)}{p(z)}. \tag{9-36}$$

En esta expresión $p(z)$ representa la probabilidad incondicional del resultado dado del experimento z, es decir, la probabilidad de que vaya a ocurrir el resultado z en un estado arbitrario de la naturaleza. Partiendo de esto, la probabilidad $p(z)$ puede presentarse del modo siguiente:

$$p(z) = p(z\,|\,\vartheta_1 \cup \vartheta_2 \cup \ldots \cup \vartheta_m). \tag{9-37}$$

Aplicando a esta expresión la fórmula de la probabilidad completa (5-42), obtenemos

$$p(z) = \sum_\vartheta p_\vartheta(z)\,\xi(\vartheta). \tag{9-38}$$

Teniendo en cuenta (9-38) la fórmula para determinar la distribución aposteriorística de probabilidades $\xi_z(\vartheta)$ adquiere el aspecto:

$$\xi_z(\vartheta) = \frac{p_\vartheta(z)\,\xi(\vartheta)}{\sum\limits_\vartheta p_\vartheta(z)\,\xi(\vartheta)}; \tag{9-39}$$

En la teoría de las probabilidades esta fórmula recibió la denominación de fórmula de Bayes.

Ejemplo 9-11. Determinemos la distribución aposterioristica de probabilidades en el problema de la línea tecnológica. Realicemos los cálculos según la fórmula (9-39) por el método tabular. El proceso del cálculo se brinda en la tabla 9-8.

Tabla 9-8

Cálculo de la distribución aposteriorística de las probabilidades en el problema de la línea tecnológica

ϑ	$\xi(\vartheta)$	$p_\vartheta(z)$			$p_\vartheta(z)\,\xi(\vartheta)$			$\xi_z(\vartheta)$		
		z_1	z_2	z_3	z_1	z_2	z_3	z_1	z_2	z_3
ϑ_1	0,6	0,60	0,25	0,15	0,36	0,15	0,09	0,818	0,555	0,310
ϑ_2	0,4	0,20	0,30	0,50	0,08	0,12	0,20	0,182	0,445	0,690
$\sum_\vartheta p_\vartheta(z)\,\xi(\vartheta)$					0,44	0,27	0,29			

Sirven de datos de partida para calcular la distribución apriorística de probabilidades $\xi(\vartheta)$, tomada de la tabla 9-2 y la distribución condicional de probabilidades $p_\vartheta(z)$ tomada de 9-5.

El valor hallado de las probabilidades aposteriorísticas demuestra que con los resultados del experimento z_1 o z_3 disminuye grandemente la indeterminación respecto a los estados de la naturaleza.

No obstante, con el resultado z_2 la indeterminación se vuelve incluso mayor que antes de realizar el experimento.

c) Principio de verosimilitud máxima

El conocimiento de las probabilidades aposteriorísticas permite estimar el estado de la naturaleza utilizando el principio de verosimilitud máxima. Según este principio se toma como estimación del estado de la naturaleza $\hat{\vartheta}$ aquel que parece el más probable basándose en los datos experimentales.

Ejemplo 9-12. Estimemos el estado de la naturaleza $\hat{\vartheta}$ en el problema del proceso tecnológico siendo z_1 el resultado del experimento. De acuerdo con la tabla 9-8 las probabilidades aposteriorísticas de los estados de la naturaleza ϑ_1 y ϑ_2 con el resultado del experimento z_1 son:

$$\xi_{z_1}(\vartheta_1) = 0{,}818, \quad \xi_{z_1}(\vartheta_2) = 0{,}182.$$

Por cuanto el máx $[\xi_{z_1}(v_1), \xi_{z_1}(v_2)] = \xi_{z_1}(\vartheta_1)$, en virtud del principio de verosimilitud máxima $\hat{\vartheta} = \vartheta_1$.

El principio de verosimilitud máxima se utiliza frecuentemente para elegir la decisión en el *problema de dos alternativas*. En este problema el espacio de estados de la naturaleza consta de dos elementos: $\Theta = \{\vartheta_1, \vartheta_2\}$ con distribución apriorística de las probabilidades $\xi(\vartheta) = (\zeta, 1 - \zeta)$. El espacio de decisiones del estadista también está compuesto de dos elementos $A = \{a_1, a_2\}$, donde a_1 es la decisión de que la naturaleza se halla en el estado ϑ_1; a_2 es la decisión de que la naturaleza se encuentra en el estado ϑ_2. Se conocen también las distribuciones condicionales de probabilidades de los resultados del experi-

mento $p_\vartheta(z)$ para $\vartheta = \vartheta_1$ y $\vartheta = \vartheta_2$. Por los resultados del experimento $z \in Z$ es necesario tomar la decisión $a \in \{a_1, a_2\}$.

Ejemplo 9-13. Problema de la estación de radiolocalización (E. R. L.). La señal puede surgir en la pantalla de la E. R. L. como resultado de la presencia del objetivo en la zona de acción de la E. R. L. o a causa de la influencia de interferencias de diversa índole. Por consiguiente, pueden tener lugar dos estados de la naturaleza: ϑ_1, hay objetivo; ϑ_2, no hay objetivo. Al mirar la pantalla puede tomarse una de las dos decisiones: a_1, hay objetivo; a_2, no hay objetivo. Con esto pueden cometerse errores de dos tipos:

$a_1|\vartheta_2$, el error del primer género denominado a veces "falsa alarma";

$a_2|\vartheta_1$, el error del segundo género llamado en ocasiones "no detección del objetivo".

Para tomar la decisión en el problema de dos alternativas se emplea con frecuencia la *relación de verosimilitud* definida por la relación

$$\Lambda(z) = \frac{\xi_z(\vartheta_1)}{\xi_z(\vartheta_2)}. \tag{9-40}$$

Se dice que tiene lugar una comprobación con respecto a la verosimilitud si existe un número determinado k tal que se toma la decisión de acuerdo con la regla siguiente:

se toma la decisión a_1 si $\Lambda(z) > k$;
se toma la decisión a_2 si $\Lambda(z) < k$;
se toma la decisión a_1 o a_2 si $\Lambda(z) = k$.

El valor de k se elige dependiendo de las consecuencias a las que puede conducir una decisión errónea. Así pues, en el problema de la E. R. L. el error de tipo "falsa alarma" puede acarrear mayores consecuencas que el error "no detección del objetivo". Sólo debe tomarse la decisión a_1 en caso en que tengamos una gran seguridad de que en realidad existe el objetivo, es decir, si se cumple la condición $\xi_z(\vartheta_1) \gg \xi_z(\vartheta_2)$. Esto significa que en el caso dado hay que tomar $k \gg 1$.

d) Determinación de la decisión de Bayes sobre la base del empleo de las probabilidades aposteriorísticas

Como vimos al principio del presente capítulo, la dificultad principal para resolver el problema con realización del experimento consiste en el brusco incremento del número de estrategias al aumentar el número de resultados posibles del experimento. Sin embargo, en realidad no hay ninguna necesidad de analizar todos los resultados del experimento posibles. Si el experimento efectuado nos proporciona algún resultado concreto $z \in Z$, entonces hay que resolver el problema para dicho resultado.

Esto puede hacerse calculando con el resultado dado del experimento z la distribución apriorística de probabilidades $\xi_z(\vartheta)$ en el espacio de estados de la naturaleza Θ. En este caso, se conocerán el espacio de estados de la naturaleza Θ, el espacio de decisiones A y la distribución de probabilidades en el espacio de estados de la naturaleza $\xi_z(\vartheta)$ en el cual está tomado en cuenta el resultado del experimento. Pero éste sólo difiere del problema sin experimento en que aquí se utiliza la distribución aposteriorística de probabilidades $\xi_z(\vartheta)$ en lugar de la apriorística $\xi(\vartheta)$. Por consiguiente, los métodos de resolución de este problema son análogos a los que se emplearon en el problema sin experimento.

Al emplear las probabilidades aposteriorísticas $\xi_z(\vartheta)$ en calidad de estrategias puras del estadista se emplean los elementos del espacio de decisiones $A = \{a_1, \ldots, a_l\}$. Con esto a cada decisión $a \in A$ van a corresponder las pérdidas del estadista $L(\xi_z, a)$ que se determinan por analogía con (9-2) como

$$L(\xi_z, a) = M_{\vartheta \mid z}[L(\vartheta, a)] = \sum_{\vartheta} L(\vartheta, a)\, \xi_z(\vartheta) \tag{9-41}$$

o teniendo en cuenta (9-39)

$$L(\xi_z a) = \frac{\sum_{\vartheta} L(\vartheta, a)\, p_{\vartheta}(z)\, \xi(\vartheta)}{\sum_{\vartheta} p_{\vartheta}(z)\, \xi(\vartheta)}. \tag{9-42}$$

El principio de Bayes se reduce a elegir tal operación $a^* \in A$ en que sea mínima la magnitud de las pérdidas determinadas según (9-42):

$$R^*(\xi_z) = R(\xi_z, a^*) = \min_{a} L(\xi_z, a). \tag{9-43}$$

e) Problema de dos alternativas

Ilustremos los principios expuestos mediante el ejemplo de la resolución del problema con distribución apriorística de las probabilidades $\xi(\vartheta = (\zeta, 1 - \zeta)$. Siendo correctas las soluciones, consideramos la pérdidas iguales a cero. El error del primer género $(a_2 | \vartheta_1)$ da una pérdida w y el del segundo género, $(a_1 | \vartheta_2)$, una pérdida de 1. Con esto, la matriz de pérdidas tiene el aspecto de la tabla 9-9.

Tabla 9-9

Pérdidas en el problema de dos alternativas

ϑ	a_1	a_2
ϑ_1	0	w
ϑ_2	1	0

Examinemos la función de decisión $d(z)$ que, según (9-27), divide el espacio de resultados del experimento Z en dominios S y $C(S)$ tales que se toma la decisión a_1, si $z \in S$ y se toma la decisión a_2, si $z \in C(S)$. Por cuanto S y $C(S)$ deben ser compactos, la resolución del problema se reduce a hallar el límtie que divide el conjunto Z en los subconjuntos no intersecados S y $C(S)$. Designemos por z_0 los elementos de $z \in Z$ que forman dicho límite. En la figura 9-7 se muestra el tipo de límite z_0 para los casos unidimensional y bidimensional.

Para hallar la ecuación que define el límite z_0, escribiremos las expresiones para las pérdidas medias estando dadas la z y las decisiones a_1 y a_2. Tomando en consideración (9-42) y los datos de

la tabla 9-9, obtenemos:

$$L(\xi_z, a_1) = \frac{p_{\vartheta_2}(z)(1-\zeta)}{\sum_{\vartheta} p_{\vartheta}(z)\,\xi(\vartheta)}; \qquad (9\text{-}44)$$

$$L(\xi_z a_2) = \frac{w p_{\vartheta_1}(z)\,\zeta}{\sum_{\vartheta} p_{\vartheta}(z)\,\xi(\vartheta)}. \qquad (9\text{-}45)$$

El límite z_0 corresponde a la igualdad de las pérdidas medias con las soluciones a_1 y a_2 lo que da:

$$(1-\zeta)\,p_{\vartheta_2}(z_0) = \zeta w p_{\vartheta_1}(z_0) \qquad (9\text{-}46)$$

o bien

$$\frac{p_{\vartheta_2}(z_0)}{p_{\vartheta_1}(z_0)} = \frac{\zeta w}{1-\zeta} = c(\zeta,\ w). \qquad (9\text{-}47)$$

Como vemos, a cada valor de ζ corresponderá su límite z_0 y, por lo tanto, los dominios correspondientes S_ζ y $C(S_\zeta)$. Por cierto,

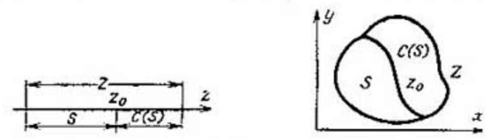


Fig. 9-7. Fraccionamiento del espacio de resultados del experimento en el problema de dos alternativas

con la magnitud ζ van a determinarse también las probabilidades de las soluciones erróneas $\alpha(\zeta)$ y $\beta(\zeta)$, que son las probabilidades de que con $\vartheta = \vartheta_1$ el punto z caerá en el dominio $C(S_\zeta)$ y se tomará la decisión a_2 y la probabilidad de que con $\vartheta = \vartheta_2$ el punto z caerá en el dominio S_ζ y se tomará la decisión a_1. Estas probabilidades se determinan por las relaciones

$$\alpha(\zeta) = P(a_2 \mid \vartheta_1) = \sum_{z \in C(S_\zeta)} p_{\vartheta_1}(z); \qquad (9\text{-}48)$$

$$\beta(\zeta) = P(a_1 \mid \vartheta_2) = \sum_{z \in S_\zeta} p_{\vartheta_2}(z). \qquad (9\text{-}49)$$

Para determinar el carácter de la dependencia de la probabilidad de soluciones erróneas de ζ, determinemos los valores $\alpha(\zeta)$ y $\beta(\zeta)$ para los valores límites $\zeta = 0$ y $\zeta = 1$.

Los elementos $z \in Z$ que forman parte del conjunto S se determinan por la condición

$$L(\xi_z, a_1) \leqslant L(\xi_z, a_2)$$

o bien

$$\frac{p_{\vartheta_2}(z)}{p_{\vartheta_1}(z)} \leqslant \frac{\zeta w}{1-\zeta}. \qquad (9\text{-}50)$$

Con $\zeta = 0$ esta condición da:

$$\frac{p_{\vartheta_2}(z)}{p_{\vartheta_1}(z)} = 0,$$

lo que es posible sólo en el caso de que $S_\zeta = \varnothing$, $C(S_\zeta) = Z$. Con esto, mediante (9-48) y (9-49) hallamos que $\alpha(0) = 1$ y $\beta(0) = 0$.

Con $\zeta = 1$ la condición (9-50) da:

$$\frac{p_{\vartheta_2}(z)}{p_{\vartheta_1}(z)} \leqslant \infty,$$

lo que determina el dominio $S_\zeta = Z$, $C(S_\zeta) = \varnothing$, de modo que $\alpha(1) = 0$, $\beta(1) = 1$.

De esta suerte, al variar ζ de 0 a 1, $\alpha(\zeta)$ varía de 1 a 0, y $\beta(\zeta)$ varía de 0 a 1.

Hallemos el riesgo de Bayes $\rho^*(\zeta)$ para la función de decisión para todos los valores posibles de ζ. Utilizando la expresión (9-29) examinada $d(z)$, con el objeto de determinar las pérdidas medias

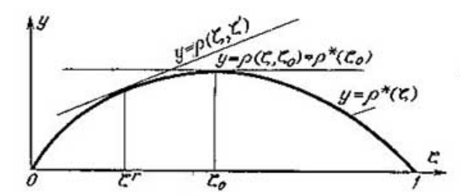


Fig. 9-8. Curva del riesgo de Bayes en el problema de dos alternativas

para la función de decisión $\rho(\vartheta, d)$ y promediando esta función en todos los estados de la naturaleza, hallamos que

$$\rho^*(\zeta) = \zeta w \alpha(\zeta) + (1 - \zeta)\beta(\zeta). \tag{9-51}$$

En la figura 9-8 se muestra la gráfica de dependencia de $\rho^*(\zeta)$ trazada según esta fórmula teniendo en cuenta el carácter hallado de la variación de $\alpha(\zeta)$ y $\beta(\zeta)$. En esta gráfica se ve que $\rho^*(\zeta)$ se convierte en cero con $\zeta = 0$ y $\zeta = 1$, y alcanza su máximo con cierta $\zeta = \zeta_0$. El valor ζ_0 determina la peor estrategia de la naturaleza para el estadista, con la cual él, utilizando el principio de Bayes, tiene las pérdidas $\rho^*(\zeta_0)$.

En la práctica, muy frecuentemente se presentan casos en que hay insuficientes datos estadísticos preliminares para determinar con exactitud la probabilidad apriorística ζ. Por eso, tiene importancia hallar la magnitud del riesgo de Bayes para el caso en que el estadista parte de cierto valor de la probabilidad apriorística ζ', y de acuerdo con esto determina las magnitudes $S_{\zeta'}$, $\alpha(\zeta')$ y $\beta(\zeta')$, mientras que el valor real de ζ será otro. En este caso el riesgo de

Bayes se determina por medio de la expresión

$$\rho(\zeta, \zeta') = \zeta w \alpha(\zeta') + (1 - \zeta)\beta(\zeta'), \qquad (9\text{-}52)$$

que es lineal respecto a ζ y define la tangente a la curva $y = \rho^*(\zeta)$ en el punto $\zeta = \zeta'$. En la figura 9-8 se ve que las pérdidas determinadas por la función $\rho(\zeta, \zeta')$, serán mayores que para la función $\rho^*(\zeta)$ si $\zeta \neq \zeta'$, de modo que el conocimiento inexacto de la distribución apriorística de probabilidades provoca el aumento de las pérdidas. Es más, si las ζ difieren mucho de ζ', las pérdidas determinadas por la función $\rho(\zeta, \zeta')$ pueden exceder de la magnitud $\rho^*(\zeta_0)$ correspondiente a la magnitud máxima de las pérdidas de Bayes. No obstante, si se parte de la distribución de probabilidades más desfavorable correspondiente al valor ζ_0, la función $y = \rho(\zeta, \zeta_0)$ va a ser continua y con cualquier ζ determina pérdidas iguales a $\rho^*(\zeta_0)$. Este caso corresponde al empleo de la estrategia de minimax.

De esta manera, únicamente es racional el empleo del principio de Bayes en los casos en que se conoce con suficiente exactitud la distribución apriorística de probabilidades $\xi(\vartheta)$. Ahora, si se desconoce o conoce de manera imprecisa la distribución apriorística de probabilidades $\xi(\vartheta)$, puede resultar más provechoso el empleo del principio de minimax.

9-5. JUEGOS ESTADISTICOS CON MUESTRAS CONSECUTIVAS

a) Observaciones preliminares

Al estudiar el juego estadístico con experimento único se señaló que éste no constaba necesariamente de una sola prueba. El puede consistir asimismo en una secuencia de pruebas, pero deben fijarse previamente la extensión y el orden de las mismas. Así pues, si se ha determinado de antemano que la decisión se toma después de que se han realizado N pruebas consecutivas, entonces toda esta secuencia de pruebas se considera como un experimento único cuyo resultado será la magnitud multidimensional $z = (z_1, \ldots, z_N)$, donde z_i, $i = 1, \ldots, N$ es el resultado de la prueba i. En virtud de esto, el juego con experimento único se denomina frecuentemente juego con extensión predeterminada de muestras entendiendo por extensión de muestra el número de experimentos realizados consecutivamente.

En los juegos con extensión de muestra predeterminada el estadista sólo puede utilizar aquellas estrategias que determinan el carácter de sus acciones después de finalizar toda la sucesión de pruebas. No obstante, existe otra posibilidad para el estadista. Así, en lugar de realizar todas las N pruebas, después de cada prueba sucesiva él puede decidir: poner fin a la prueba y elegir de entre A cierta decisión, fundándose en la información ya existente,

o realizar la prueba siguiente. Esto amplía la clase de las estrategias posibles del estadista ya que a la elección de la decisión de entre el conjunto A se añade además la elección de la decisión de terminar o continuar el experimento. Tales juegos se denominan *juegos con muestra consecutiva.* Por cierto, si está determinado un número límite admisible de pruebas después de cuya realización deberá tomarse obligatoriamente una decisión de entre A, el juego se llama de *muestra consecutiva truncada.* Limitaremos la exposición ulterior sólo a estos juegos remitiendo a los interesados en un estudio más detallado de la materia a la literatura a propósito. Vamos a denominar *observaciones* a los resultados consecutivos separados del experimento.

Si el experimento no costara nada, no tendría sentido ampliar la clase de estrategias introduciendo muestras consecutivas ya que el estadista no perdería nada y al efectuar todas las N observaciones sólo podría ganar. Pero en muchos casos el experimento es costoso y requiere pérdidas de tiempo. Entonces el estadista puede obtener una considerable ganancia si en cada etapa del experimento compara el costo de continuación del mismo con la ganancia esperada debido a la información adicional obtenida.

Examinemos el método de descripción del juego con muestras consecutivas truncadas. Designemos por Z_i el conjunto de resultados de la prueba i. Entonces el espacio completo de resultados del experimento Z al efectuar todas las N pruebas puede representarse en la forma

$$Z = Z_1 \times \ldots \times Z_N. \quad (9\text{-}53)$$

En el juego con muestras consecutivas se admite tomar la decisión basándose en la realización no de todas las N observaciones $z_1, \ldots, z_N$, sino únicamente las primeras j de ellas $z_1, \ldots, z_j$. En virtud de esto puede fraccionarse el conjunto Z en subconjuntos no intersecados $S_0, S_1, \ldots, S_N$ tales que la decisión se toma fundándose en la realización de las primeras j observaciones si $z \in S_j$.

El conjunto $S = (S_0, S_1, \ldots, S_N)$ se denomina *plan de muestra consecutiva.* El conjunto Z puede dividirse por varios métodos diferentes en subconjuntos no intersecados S_j. Cada uno de estos métodos determinará su plan de muestra consecutiva. Designaremos por $\mathfrak{C}$ y llamaremos *clase completa de muestras consecutivas* al conjunto de diversos planes de muestra consecutiva posibles.

Para que sea factible la toma de la decisión, hay que presentar la función de decisión $d(z)$ que determina la decisión de A para cada secuencia de observaciones. Al igual que en el juego con experimento único, el estadista elige la función de decisión $d(z)$ de entre el espacio de funciones de decisión D del que forman parte todas las funciones de decisión posibles.

La estrategia del estadista en el juego con muestras consecutivas consiste, en primera, en la elección del plan de decisión consecutiva $S \in \mathfrak{C}$ que indica cuándo debe terminarse el experi-

mento, y en segunda, en la elección de la función de decisión $d \in D$ la cual señala qué decisión debe tomarse al finalizar el experimento. De este modo el par (S, d) determina la estrategia del estadista. Tomando en consideración que el par (S, d) es un elemento del producto directo $\mathfrak{S} \times D$, llegamos a la conclusión de que $\mathfrak{S} \times D$ es el espacio de estrategias puras del estadista.

El número total de estrategias en los juegos con muestras consecutivas resulta mucho mayor que en los juegos con experiencia única lo que provoca grandes dificultades al redactar la lista completa de estrategias del estadista y elegir la mejor de ellas. Este trabajo puede simplificarse considerablemente si en cada etapa del experimento se calcula la distribución aposteriorística de probabilidades utilizando para ello toda la información acumulada hasta el momento dado. A medida que se efectúan las observaciones consecutivas, disminuye la indeterminación respecto al verdadero estado de la naturaleza ya que surge la posibilidad de elaborar un criterio para determinar el momento de dar fin al experimento y tomar la decisión de entre el conjunto A.

b) Utilización de la distribución aposteriorística de probabilidades para determinar las reglas consecutivas de Bayes

Solamente examinaremos aquellos juegos estadísticos en los que los resultados de algunas subpruebas $z_1, \ldots, z_N$ son magnitudes aleatorias independientes. Designemos por $q_\vartheta(z_j)$ la distribución de probabilidades en el espacio Z_j con el estado de la naturaleza dado ϑ. Vamos a considerar constante e igual a $c = 1$ el valor de la subprueba aislada.

En calidad de fundamento para escribir el plan de muestras consecutivas tomaremos la distribución de probabilidades $\xi(\vartheta)$ en el espacio de estados de la naturaleza Θ, al hacerlo, por distribución $\xi(\vartheta)$ vamos a entender no sólo la distribución apriorística de probabilidades antes de iniciar el experimento, sino también la distribución *aposteriorística* de éstas después de realizar varias subpruebas. En lo sucesivo vamos a designar por $\xi_0(\vartheta)$, la distribución apriorística de probabilidades y por $\xi_j(\vartheta)$, la distribución aposteriorística de éstas después de realizar j subpruebas.

La distribución de probabilidades $\xi_j(\vartheta)$ contendrá en sí toda la información sobre el estado de la naturaleza ϑ que se tenía antes de realizar el experimento y también la que se obtuvo a resultas de efectuar las primeras j observaciones. Por eso, la distribución $\xi_j(\vartheta)$ puede considerarse como apriorística antes de la observación $(j+1)$. Luego entonces, la distribución $\xi_{j+1}(\vartheta)$ se expresa por $\xi_j(\vartheta)$ con ayuda de la fórmula para la distribución aposteriorística de probabilidades. Vamos a considerar el paso de la distribución apriorística de probabilidades $\xi_j(\vartheta)$ a la aposteriorística $\xi_{j+1}(\vartheta)$ como cierta transformación T de la distribución $\xi_j(\vartheta)$. En-

tonces la distribución aposterioristica de probabilidades puede escribirse según (9-39) en la forma

$$T\xi_j(\vartheta) = \xi_{j+1}(\vartheta) = \frac{\xi_j(\vartheta)\, q_\vartheta(z_{j+1})}{\sum_\vartheta \xi_j(\vartheta)\, q_\vartheta(z_{j+1})}. \tag{9-54}$$

Analicemos el principio para obtener el plan de muestra consecutiva basado en el empleo de la distribución aposteriorística de probabilidades, mediante el ejemplo del problema de dos alternativas en el cual $\Theta = \{\vartheta_1 \vartheta_2\}$ y $A = \{a_1, a_2\}$.

Designemos por $(\zeta, 1 - \zeta)$ la distribución aposteriorística de probabilidades en el espacio Θ después de realizar varias subpruebas. En este problema el espacio E de estrategias mixtas de la naturaleza se determina por el dominio de valores posibles ζ, o sea, el intervalo [0, 1] del eje real.

Decisión sobre la continuación del experimento
Decisión a_2 · Decision a_1 · ζ
0 · $\Delta(a_2)$ · γ · δ · $\Delta(a_1)$ · 1

Fig. 9-9. Regla de la toma de decisión en el problema de dos alternativas

Puede resultar que $\zeta = 1$. En este caso sólo es posible tomar la decisión a_1. Si $\zeta = 0$ se toma sin falta la decisión a_2. Pero si, por ejemplo, resulta que $\zeta = 0{,}5$, entonces es imposible dar preferencia a decisión alguna y hay que continuar el experimento para precisar el verdadero estado de la naturaleza.

Aqui se examinaron casos extremos de distribución de probabilidades $\xi(\vartheta)$. En el caso general pueden establecerse magnitudes δ y γ $(0 \leqslant \delta \leqslant 1,\ 0 \leqslant \gamma \leqslant 1,\ \delta \geqslant \gamma)$ tales que:

si ζ está en la gama $[\delta, 1]$ se toma la decisión a_1;

si ζ está en la gama $[0, \gamma]$ se toma la decisión a_2;

si ζ está en la gama $[\gamma, \delta]$ se toma la decisión de realizar la subprueba siguiente. Las gamas $\Delta(a_1) = [\delta, 1]$ y $\Delta(a_2) = [0, \gamma]$ se llaman *dominios de parada*. La regla para tomar la decisión está representada gráficamente en la figura 9-9.

Los dominios de parada pueden determinarse también para el caso en que el número de estados de la naturaleza es mayor de dos. Aunque es verdad que entonces el espacio de estrategias mixtas de la naturaleza E tendrá una forma más compleja. Así pues, en el caso de tres estados de la naturaleza tiene la forma de triángulo equilátero con altura igual a la unidad (véase la fig. 9-11).

En el caso general se llama dominio de parada $\Delta(a)$ en el espacio E subconjunto de este espacio

$$\Delta(a) \subset \mathrm{E} \tag{9-55}$$

tal que si después de cierta subprueba resulta que $\xi(\vartheta) \in \Delta(a)$, entonces se pone fin al experimento y se toma la decisión a,

Por cuanto cada subprueba tiene su valor, no será indiferente si $\xi(\vartheta)$ cae en la zona $\Delta(a)$ antes de realizar el experimento o después de varias subpruebas. Esto significa que con cada nueva subprueba van a cambiar tanto los valores de $\xi(\vartheta)$ como los dominios $\Delta(a)$. Por cierto, en cada etapa del experimento presentan interés no aquellas subpruebas que ya se efectuaron y cuyos resultados se utilizaron ya, sino las que quedan por realizar hasta el final del experimento. En concordancia con esto designaremos por $\Delta_{N-j}(a)$ los dominios de parada en la etapa del experimento en que se llevaron a cabo j subpruebas de N.

Si el espacio de soluciones consta de m elementos $a_1, \ldots, a_m$, en cada etapa del experimento deben determinarse m dominios $\Delta(a)$ correspondientes a cada $a \in A$. Estos dominios deben ser convexos y no intersecados.

c) Regla de las muestras consecutivas

Supongamos que en el espacio E se han separado m dominios $\Delta(a)$ en cada etapa del experimento. Entonces, la regla de las muestras consecutivas va a consistir en lo siguiente.

Inicialmente se conoce la distribución apriorística de probabilidades $\xi_0(\vartheta)$. Si $\xi_0(\vartheta) \in \Delta_N(a_i)$ para alguna i se toma la decisión a_i sin realizar experimento. Si $\xi_0(\vartheta) \notin \Delta_N(a_i)$ con ninguna i, se realiza la primera subprueba y se calcula la distribución aposteriorística de probabilidades $\xi_1(\vartheta)$. Después, se examinan los dominios $\Delta_{N-1}(a)$. Si $\xi_1(\vartheta) \in \Delta_{N-1}(a_i)$ para alguna i, se pone fin al experimento y se toma la decisión a_i. Si $\xi_1(\vartheta) \notin \Delta_{N-1}(a_i)$ para ninguna i, se realiza la siguiente subprueba, etc.

Por lo general, si no se dio por terminado el experimento durante las primeras $j - 1$ observaciones, se realiza una observación complementaria y se calcula $\xi_j(\vartheta)$. Si $\xi_j(\vartheta) \in \Delta_{N-j}(a_i)$ para cualquier i, entonces se termina el experimento y se toma la decisión a_i. Si la elección no se realizó después de la observación $(N - 1)$ se efectúa la observación N y se realiza la elección definitiva. Para que se cumpla esta condición los dominios $\Delta_0(a_i)$ deben elegirse de tal suerte que

$$\bigcup_i \Delta_0(a_i) = \mathrm{E}. \qquad (9\text{-}56)$$

Como vemos, el problema de describir las reglas consecutivas de Bayes se reduce a la tarea de determinar los dominios $\Delta_{N-j}(a)$ en el espacio E para todas las $a \in A$ y para todas las j de 0 a N.

d) Función de riesgo con la regla consecutiva óptima

La función de riesgo con muestras consecutivas debe determinar en cada etapa del experimento (por ejemplo, después de j subpruebas), las pérdidas medias mínimas que sufrirá el estadista

tomando la mejor de todas las decisiones posibles incluyendo la de que el experimento debe continuarse. Por cuanto la toma de decisión se basa en los conocimientos de la distribución aposteriorística de probabilidades $\xi_j(\vartheta)$ la función de riesgo dependerá también de esta distribución. Designemos por $\rho^*(\xi_j)$ la función de riesgo que se obtiene después de realizar j subpruebas.

Busquemos consecutivamente, partiendo de la última etapa del experimento, las expresiones que definen la función de riesgo. Supongamos que se han realizado todas las N subpruebas y está hallada la distribución apriorística de probabilidades $\xi_N(\vartheta)$. Las pérdidas medias que sufre con esto el estadista al tomar la decisión $a \in A$ son:

$$L(\xi_N, a) = \sum_{\vartheta} L(\vartheta, a)\, \xi_N(\vartheta). \tag{9-57}$$

El valor mínimo de estas pérdidas correspondiente a la solución de Bayes a^* se determina por la expresión

$$R^*(\xi_N) = L(\xi_N, a^*) = \min_a L(\xi_N, a). \tag{9-58}$$

Por cuanto no es posible tomar ninguna decisión mejor que a^*, la magnitud $R^*(\xi_N)$ va a coincidir en esta etapa del experimento con la función de riesgo

$$\rho^*(\xi_N) = R^*(\xi_N). \tag{9-59}$$

Examinemos ahora una etapa arbitraria del experimento y hallemos las pérdidas medias mínimas que sufrirá el estadista después que realizó j subpruebas ($j = 0, 1, \ldots, N-1$). En este caso él dispone de la distribución de probabilidades $\xi_j(\vartheta)$.

En el caso dado el estadista puede obrar de dos maneras.

1. Puede poner fin al experimento y tomar la decisión $a \in A$. Las pérdidas medias mínimas que él sufre son:

$$R^*(\xi_j) \min_a L(\xi_j, a) = \min_a \sum_{\vartheta} L(\vartheta, a)\, \xi_j(\vartheta). \tag{9-60}$$

2. Puede tomar la decisión de realizar la subprueba $(j+1)$. En tal caso va a sufrir, en primera, las pérdidas iguales al costo de la prueba las cuales tomamos como unidad, y, en segunda, las pérdidas que sufre después de tomar la mejor decisión por los resultados de la observación $(j+1)$, es decir, $\rho^*(\xi_{j+1})$.

No obstante, el estadista no sabe la distribución de probabilidades $\xi_{j+1}(\vartheta)$. Conoce sólo la distribución $\xi_j(\vartheta)$ y únicamente puede determinar $\xi_{j+1}(\vartheta)$ por la fórmula (9-54) como $T\xi_j(\vartheta)$. Con esto, la magnitud $T\xi_j(\vartheta)$ puede ser hallada para los valores concretos de $z_{j+1} \in Z_{j+1}$ y $\vartheta \in \Theta$ que el estadista no conoce por ahora. Por esto no debe referirse a un valor concreto de la función de riesgo $\rho^*(\xi_{j+1})$, sino únicamente al valor medio $M[\rho^*(T\xi_j)]$, por cierto, el promedio debe tomarse por todos los valores posibles de

$z_{j+1} \in Z_{j+1}$ con la distribución de probabilidades $q_\vartheta(z_{j+1})$ y por todas las $\vartheta \in \Theta$ con la distribución de probabilidades $\xi_j(\vartheta)$:

$$M[\rho^*(T\xi_j)] = \sum_{\vartheta, z_{j+1}} \rho^*[T\xi_j(\vartheta)]\, q_\vartheta(z_{j+1})\, \xi_j(\vartheta). \tag{9-61}$$

De este modo, las pérdidas mínimas del estadista en caso de que decida continuar el experimento, son:

$$1 + M[\rho^*(T\xi_j)]. \tag{9-62}$$

Después de realizar j subpruebas la función de riesgo se determina por el mínimo de pérdidas medias al examinar ambos métodos de operación del estadista: terminación del experimento y continuación del mismo. Por tanto,

$$\rho^*(\xi_j) = \text{mín}\,\{R^*(\xi_j),\ 1 + M[\rho^*(T\xi_j)]\}. \tag{9-63}$$

La expresión (9-63) puede escribirse de modo más compacto sin emplear el subíndice j que designa el número de la subprueba. Vamos a examinar la etapa del experimento en que hasta su terminación quedan k subpruebas. Designemos por $\xi(\vartheta)$ la distribución aposteriorística de probabilidades en esta etapa y por $\xi'(\vartheta) = T\xi(\vartheta)$ la distribución aposteriorística de probabilidades después de realizar una subprueba más. Designemos por $\rho_k^*(\xi)$ la función de riesgo en la fase cuando hasta el final del experimento quedan k subpruebas. Entonces las expresiones (9-59) y (9-63) se escriben en la forma

$$\rho_0^*(\xi) = R^*(\xi); \tag{9-64}$$

$$\rho_k^*(\xi) = \text{mín}\left\{R^*(\xi),\ 1 + M\left[\rho_{k-1}^*(T\xi)\right]\right\}. \tag{9-65}$$

Designemos por V el espacio de resultados de la subprueba siguiente cuyos elementos vamos a designar por v. Con estas designaciones obtenemos:

$$R^*(\xi) = \min_a \sum_\vartheta L(\vartheta, a)\, \xi(\vartheta); \tag{9-66}$$

$$M\left[\rho_{k-1}^*(T\xi)\right] = \sum_{\vartheta, v} \rho_{k-1}^*[T\xi(\vartheta)]\, q_\vartheta(v)\, \xi(\vartheta); \tag{9-67}$$

$$T\xi(\vartheta) = \frac{\xi(\vartheta)\, q_\vartheta(v)}{\sum_\vartheta \xi(\vartheta)\, q_\vartheta(v)}. \tag{9-68}$$

Las fórmulas (9-64) y (9-65) pueden utilizarse para determinar los dominios de parada. Sin embargo, sólo se obtienen resultados relativamente sencillos para el problema de dos alternativas a cuyo estudio nos limitaremos en lo adelante.

e) Determinación de los dominios de parada para el problema de dos alternativas con muestra consecutiva truncada

Examinemos el caso del problema de dos alternativas con función de pérdidas determinada por la matriz

$$\mathbf{Q}=\begin{Vmatrix} 0 & q_{12} \\ q_{21} & 0 \end{Vmatrix}, \qquad (9\text{-}69)$$

y con la distribución de probabilidades en el espacio de estado de la naturaleza $\xi(\vartheta)=(\zeta,\ 1-\zeta)$. Designemos por $\Delta_k(a_1)$ y $\Delta_k(a_2)$ los dominios de parada correspondientes a las soluciones a_1 y a_2 cuando hasta el final del experimento quedan por realizar k observaciones. Estos dominios se determinan por los valores $\zeta=\delta_k$ y $\zeta=\gamma_k$, $\gamma_k\leqslant\delta_k$ tales que

$$\Delta_k(a_1)=[\delta_k\leqslant\zeta\leqslant 1],\qquad \Delta_k(a_2)=[0\leqslant\zeta\leqslant\gamma_k]. \qquad (9\text{-}70)$$

La determinación de las zonas de parada se reduce a determinar los valores de γ_k y δ_k para $k=0,\ 1,\ \ldots,\ N$. Esto puede hacerse utilizando las expresiones obtenidas para la función de riesgo.

Iniciemos la determinación de los puntos límites de los dominios de parada partiendo de $k=0$. Al expresar $\xi(\vartheta)$ por ζ, se obtiene de (9-64) y (9-66):

$$\rho_0^*(\zeta)=R^*(\zeta)=\min\left[(1-\zeta)\,q_{21},\ \zeta q_{12}\right]. \qquad (9\text{-}71)$$

Según (9-56), $\Delta_0(a_1)$ y $\Delta_0(a_2)$ deben ser conjuntos reales no intersecados cuya reunión proporciona el segmento cerrado 0,1. Por consiguiente, estos dominios deben estar separados uno del otro por el punto $\xi=\gamma_0=\delta_0$ que se determina por la condición

$$(1-\zeta)\,q_{21}=\zeta q_{12} \qquad (9\text{-}72)$$

La condición (9-72) es la condición de igualdad de las pérdidas medias mínimas al tomar las decisiones a_1 y a_2. De (9-72) hallamos:

$$\gamma_0=\delta_0=\frac{q_{21}}{q_{12}+q_{21}}. \qquad (9\text{-}73)$$

Para $k\neq 0$ arbitraria la función de riesgo tiene la forma:

$$\rho_k^*(\zeta)=\min\left\{R^*(\zeta),\ 1+M\left[\rho_{k-1}^*(T\zeta)\right]\right\}; \qquad (9\text{-}74)$$

Las condiciones de obtención de los valores límites ζ para los conjuntos $\Delta_k(a_1)$ y $\Delta_k(a_2)$ serán las condiciones de igualdad de las pérdidas medias mínimas al tomar la decisión a_1 o a_2 y al decidir continuar el experimento, lo que se expresa mediante la relación

$$R^*(\zeta)=1+M\left[\rho_{k-1}^*(T\zeta)\right]. \qquad (9\text{-}75)$$

Observando que en la expresión (9-71) para $R^*(\zeta)$ la magnitud ζq_{21} es la función monótonamente creciente ζ, que se vuelve cero para $\zeta = 0$ y $(1-\zeta) q_{12}$ es la función monótonamente decreciente ζ que se convierte en cero con $\zeta = 1$, y sustituyendo en los puntos límites ζ por γ_k para el dominio $\Delta_k(a_2)$, y por δ_k para el dominio $\Delta_k(a_1)$, podemos escribir (9-75) en la forma

$$\gamma_k q_{21} = 1 + M[\rho^*_{k-1}(T\gamma_k)]; \tag{9-76}$$

$$(1-\delta_k) q_{12} = 1 + M[\rho^*_{k-1}(T\delta_k)]. \tag{9-77}$$

Estas condiciones se representan geométricamente en la figura 9-10.

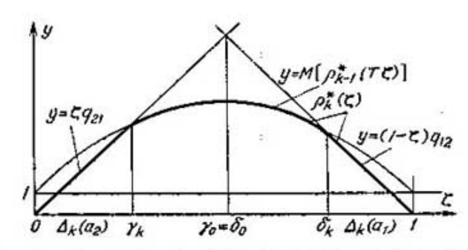


Fig. 9-10. Determinación de los dominios de parada en el problema de dos alternativas

Las relaciones obtenidas muestran que el problema de determinar γ_k y δ_k así como las funciones de riesgo $\rho^*_k(\zeta)$ se reducen a calcular la magnitud $M[\rho^*_{k-1}(T\zeta)]$. La fórmula (9-67) da la expresión general para esta magnitud con una Θ arbitraria. Para el problema de dos alternativas esta expresión adquiere la forma:

$$M[\rho^*_{k-1}(T\zeta)] = \zeta \sum_v \rho^*_{k-1}(T\zeta)\, q_{\vartheta_1}(v) + $$
$$+ (1-\zeta) \sum_v \rho^*_{k-1}[T(1-\zeta)]\, q_{\vartheta_2}(v), \tag{9-78}$$

donde

$$T\zeta = \frac{\zeta q_{\vartheta_1}(v)}{\zeta q_{\vartheta_1}(v) + (1-\zeta) q_{\vartheta_2}(v)}; \tag{9-79}$$

$$T(1-\zeta) = \frac{(1-\zeta) q_{\vartheta_2}(v)}{\zeta q_{\vartheta_1}(v) + (1-\zeta) q_{\vartheta_2}(v)}. \tag{9-80}$$

Para $k = 1$ la función $\rho^*_0(\zeta)$ determinada por la expresión (9-71) es una función muy sencilla de ζ, así que la magnitud $M[\rho^*_0(T\zeta)]$ puede determinarse directamente mediante (9-78). Conociendo $M[\rho^*_0(T\zeta)]$ puede calcularse $\rho^*_1(\zeta)$ por (9-74), luego determinarse $M[\rho^*_1(T\zeta)]$ por (9-78) y de aquí, $\rho^*_2(\zeta)$ mediante (9-74),

etc. Aunque este proceso de búsqueda de $\rho_k^{\bullet}(\zeta)$ requiere grandes cálculos, sin embargo, en este caso no es necesario efectuar operaciones más complicadas que la búsqueda de las esperanzas matemáticas.

PROBLEMAS PARA EL CAPITULO 9

9-1. Hallar, por los datos de la tabla 9-3, la estrategia fundada en el principio de minimax de las pérdidas complementarias en el problema de la línea tecnológica.

9-2. Realizar los cálculos necesarios para determinar las estrategias de minimax y Bayes en la línea tecnológica con experimento único, empleando la construcción geométrica en la figura 9-6.

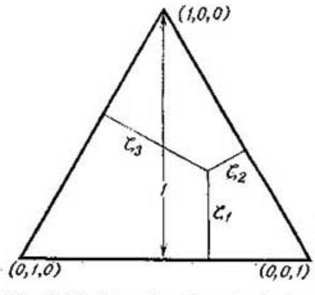


Fig. 9-11. Espacio de estrategias mixtas de la naturaleza para el caso de tres estados de la naturaleza

9-3. Por un canal de comunicación se transmite la señal u sobre la cual está superpuesta la interferencia s distribuida por la ley normal $N(0, \sigma^2)$ de suerte que en la parte de recepción se mide la magnitud $z = u + s$. La señal u puede tomar dos valores: $u = 0$, no hay señal (ϑ_1) y $u = u_0 = \text{const}$, hay señal (ϑ_2). El dispositivo de recepción tiene el umbral z_0 y suministra la decisión de que se ha transmitido la señal u_0, si $z > z_0$. Determinar la magnitud del umbral z_0 utilizando el método de resolución del problema de dos alternativas haciendo $\zeta = 0{,}25$ y $w = 2$.

9-4. Demostrar que en el caso de tres estados de la naturaleza $\Theta = \{\vartheta_1, \vartheta_2, \vartheta_3\}$ todas las estrategias mixtas del estadista $\xi(\vartheta) = (\zeta_1, \zeta_2, \zeta_3)$ forman el espacio de estrategias mixtas que tiene forma de triángulo con altura igual a la unidad como se muestra en la figura 9-11.

Capítulo décimo

PROGRAMACIÓN DINÁMICA

10-1. CONTROL OPTIMO COMO PROBLEMA VARIACIONAL

a) Enunciado matemático del problema de control óptimo

La tendencia característica en la construcción de los sistemas de control automático modernos es la aspiración de lograr sistemas que sean los mejores en determinado sentido. En el caso de control de procesos tecnológicos esta tendencia se pone de manifiesto en obtener una cantidad máxima de productos de alta calidad con empleo limitado de recursos (materias primas, energía, etc.). En los sistemas de control de naves, aviones y cohetes se tiende a minimizar el tiempo después de cuyo transcurso el objeto sale al punto o trayectoria deseados, limitándose el ángulo de desviación de los timones, la cantidad de combustible consumido, etc. En los sistemas de seguimiento y estabilización tiene importancia alcanzar una precisión máxima existiendo todas las limitaciones posibles impuestas a las coordenadas del objeto controlado, los elementos ejecutivos y el regulador. En todos estos ejemplos los problemas de control se reducen a la búsqueda del mejor proceso en determinado sentido de la palabra, de entre el conjunto de los procesos posibles, es decir, pertenecen a la clase de los problemas dinámicos de control.

Como se mostró en el capítulo 6, el enunciado matemático de los problemas dinámicos de control automático se reduce a lo siguiente. Hay un objeto controlado cuyo estado se caracteriza por la variable multidimensional $x = (x^{(1)}, \ldots, x^{(N)})$. Puede cambiarse el carácter de los procesos en el objeto controlado utilizando uno u otro procedimiento de control u del espacio de procedimientos de control admisibles U. En el caso general el control $u \in U$ puede ser asimismo la magnitud multidimensional $u = (u^{(1)}, \ldots, u^{(n)})$.

El carácter del movimiento del objeto controlado se describe por medio del sistema de ecuaciones diferenciales

$$\dot{x} = g(x, u), \quad x(0) = c. \tag{10-1}$$

Como criterio de calidad de control se emplea la estimación por la integral de la forma

$$J_1(u) = \int_0^T Q_1[x(t), u(t)]\, dt, \tag{10-2}$$

que tiene el sentido físico de pérdidas, donde T es el tiempo en que transcurre el proceso controlado y $Q_1[x(t), u(t)] = q_1(t)$ son las pérdidas instantáneas en el momento t con el estado del sistema $x(t)$ y el control $u(t)$. Pueden ser limitaciones adicionales impuestas a la cantidad de recursos o los límites de variación de algunos parámetros que se expresan mediante la relación matemática

$$\int_0^T H[x(t), u(t)]\,dt \leqslant K, \tag{10-3}$$

que coincide con (6-33).

Como quedó establecido en el capítulo 6, se denomina óptimo aquel control u del conjunto de controles admisibles U con el que para el objeto descrito por la ecuación diferencial (10-1) y las limitaciones impuestas a los recursos utilizados (10-3), el criterio de calidad de control (10-2) adquiere el valor mínimo (máximo).

El problema de control óptimo formulado de este modo pertenece a la categoría de los problemas variacionales de cuya resolución se ocupa la rama de las matemáticas que recibió la denominacion de *cálculo variacional*. La magnitud $J_1(u)$ determinada por la relación (10-2) recibió el nombre de *funcional*. A diferencia de una función, por ejemplo, $f(x)$ cuyos valores numéricos se presentan en el conjunto de valores del argumento x, los valores numéricos de la funcional $J_1(u)$ se presentan en el conjunto de todos los controles posibles $u(t)$. El problema de la búsqueda del control óptimo se reduce a elegir de entre el conjunto de controles admisibles $u(t)$ aquel con el que la funcional $J_1(u)$ adquiere el valor numérico mínimo.

Corrientemente, los problemas que requieren la minimización de la funcional del tipo (10-2) subordinada a la relación de diferencial (10-1), al existir la limitación de integral (10-3), se sustituyen por la minimización de la nueva funcional

$$J(u) = \int_0^T Q_1(x, u)\,dt + \lambda \int_0^T H(x, u)\,dt, \tag{10-4}$$

subordinada únicamente a la relación diferencial (10-1). El parámetro λ de la funcional (10-4), que recibió el nombre de factor de Lagrange, representa en los problemas de optimización de control el papel del "precio" de los recursos limitados. Su valor se halla por las condiciones límites del problema variacional.

La posibilidad de simplificar el problema variacional con las limitaciones de integral mediante la introducción de los factores de Lagrange se desprende del siguiente teorema.

Teorema 10-1. *Si $u(t)$ es el control óptimo con el cual la funcional (10-4) alcanza el mínimo absoluto y se cumple la limitación (10-3), entonces con $u(t)$ se alcanza el mínimo absoluto de la funcional (10-2) subordinada a la limitación (10-3).*

La *demostración* es por el absurdo. Sea $v(t)$ otro control distinto de $u(t)$, por cierto, tal que

$$\int_0^T Q_1(x, v)\,dt < \int_0^T Q_1(x, u)\,dt \tag{10-5}$$

y se cumple la condición

$$\int_0^T H(x, v)\,dt \leqslant K. \tag{10-6}$$

Entonces

$$\int_0^T Q_1(x, v)\,dt + \lambda \int_0^T H(x, v)\,dt \leqslant \int_0^T Q_1(x, v)\,dt + \lambda K <$$
$$< \int_0^T Q_1(x, u)\,dt + \lambda K = \int_0^T Q_1(x, u)\,dt + \lambda \int_0^T H(x, u)\,dt, \tag{10-7}$$

lo que contradice la suposición de que $u(t)$ minimiza (10-4).

La aplicación de los métodos de cálculo variacional al problema de la búsqueda del control óptimo no se difundió a causa de una serie de dificultades. Por lo tanto, no vamos a detenernos en los métodos de resolución del problema variacional remitiendo a los interesados en esta cuestión a la literatura correspondiente. Señalaremos únicamente algunas cuestiones de importancia para la exposición ulterior.

Un importantísimo concepto de cálculo variacional es el *de variación de función* que durante la investigación de las funcionales juega el mismo papel que la diferencial al investigar las funciones.

Sea $f(x)$ una función continua en el intervalo $[a, b]$. Examinemos el punto interior x de este intervalo y cierto valor fijo de la diferencial del argumento de la función $\Delta x = dx$. La diferencia

$$f(x + \Delta x) - f(x) = df(x) = f'(x)\,\Delta x \tag{10-8}$$

se llama diferencial de la función $f(x)$ en el punto x. Como se sabe, la condición $df(x) = 0$ es la condición necesaria de mínimo (máximo) de la función $f(x)$ en el punto x.

Para lograr analogía con el cálculo variacional es conveniente determinar la diferencial $df(x)$ de manera algo distinta. Examinemos el valor de la función $f(x + \varepsilon\Delta x)$; para x y Δx fijadas éste será función de ε. Diferenciando esta función por ε y haciendo $\varepsilon = 0$, obtenemos:

$$\frac{\partial}{\partial \varepsilon} f(x + \varepsilon\Delta x)\Big|_{\varepsilon=0} = f'(x)\,\Delta x = df(x). \tag{10-9}$$

Examinemos los conceptos análogos de cálculo variacional. Sean $u(t)$ y $u_1(t)$ dos controles utilizados. La diferencia

$$\delta u(t) = u_1(t) - u(t) \tag{10-10}$$

se denomina variación de la función $u(t)$ y la diferencia

$$\delta J(u) = J(u + \delta u) - J(u) \tag{10-11}$$

se llama *variación de la funcional.*

La variación de la funcional puede definirse de otro modo. Para esto examinemos con $u(t)$ y $\delta u(t)$ fijadas la funcional

$$J(u + \varepsilon\delta u) = \varphi(\varepsilon), \tag{10-12}$$

que es la función de ε. Supongamos que dicha funcional está determinada para diferentes ε, lo cual significa que son posibles distintos controles $u(t) + \varepsilon u(t)$ en las cercanías del control fijado $u(t)$, o sea, $u(t)$ es un punto interior del espacio de controles admisibles U. Entonces, la variación de la funcional puede definirse por analogía con la diferencial de la función que está dada por la relación (10-9) como

$$\delta J(u) = \frac{\partial}{\partial\varepsilon} J(u + \varepsilon\delta u)\Big|_{\varepsilon=0}. \tag{10-13}$$

Si $u(t)$ es el control óptimo, la función $\varphi(\varepsilon)$ va a alcanzar su mínimo con $\varepsilon = 0$. En este caso

$$\frac{\partial\varphi(\varepsilon)}{\partial\varepsilon}\Big|_{\varepsilon=0} = \frac{\partial}{\partial\varepsilon} J(u + \varepsilon\delta u)\Big|_{\varepsilon=0} = 0, \tag{10-14}$$

es decir, $\delta J(u) = 0$. Así pues, el control óptimo será aquel con el que la variación de la funcional se vuelve cero.

La condición $\delta J(u) = 0$ se utiliza en el cálculo variacional para obtener la llamada ecuación diferencial de Euler entre cuyo conjunto de soluciones se halla después el control $u(t)$ que minimiza la funcional (10-4).

b) Dificultades relacionadas con la resolución del problema variacional

Al buscar el mando óptimo por métodos variacionales nos enfrentamos con una serie de dificultades algunas de las cuales tienen carácter de principio:

1) los métodos variacionales sólo dan la posibilidad de hallar los máximos y mínimos relativos de la funcional $J(u)$ mientras que nos interesa encontrar el máximo o mínimo absolutos;

2) para muchos problemas técnicos las ecuaciones de Euler no resultan lineales lo que frecuentemente imposibilita obtener la solución del problema variacional en forma explícita;

3) por lo común, los valores de las señales de control llevan impuestas limitaciones que hacen imposible la búsqueda del control óptimo por los métodos variacionales.

Por cuanto esta última circunstancia tuvo una importancia decisiva para el desarrollo de nuevas ideas en la rama del control óptimo, nos detendremos en ella más detalladamente.

Las limitaciones que se imponen habitualmente a las señales de mando son del tipo

$$|u_i(t)| \leq M_i, \tag{10-15}$$

lo que significa la necesidad de limitar la magnitud de las señales suministradas a los órganos de mando. Así, están limitados la tensión máxima suministrada al inducido de un motor eléctrico, el ángulo de rotación máximo del timón de un avión, la temperatura máxima en la cámara de combustión de un motor a reacción, etc. Por cierto, el logro de procesos óptimos requiere por lo general que se mantengan las señales de mando en sus valores extremos, lo que corresponde al transcurso más rápido y efectivo de los procesos en el objeto controlado. El carácter del cambio de control

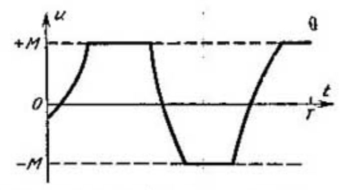


Fig. 10-1. Aspecto característico de la señal de mando óptima

$u(t)$ típico para estos casos con proceso óptimo se muestra en la figura 10-1.

Sin embargo, los valores límites del control $u(t)$ se encuentran en los límites del dominio de controles admisibles U y, por ende, no son puntos interiores de este dominio para los cuales solamente son útiles los métodos variacionales. Verdad que es posible librarse de las limitaciones del tipo (10-15) introduciendo nuevas variables v_i relacionadas con las variables u_i mediante la correlación $u_i = M_i \operatorname{sen} v_i$. Con esto, a los valores $|u_i| = M_i$ van a corresponder $v_i = \pm\pi/2$ que son puntos interiores del dominio de nuevos controles admisibles. No obstante, este cambio de variables conduce, por lo común, a considerables complicaciones de las ecuaciones obtenidas.

Las dificultades existentes en el camino de la resolución del problema de optimización del control provocaron un intenso estudio del problema de los procesos óptimos por diversos científicos en la URSS y los EE. UU. Los matemáticos soviéticos L. S. Pontriagin y sus discípulos V. G. Boltianski, R. V. Gamkrelidze y E. F. Mischenko crearon la teoría del control óptimo que se basa en el *principio del máximo* formulado por L. S. Pontriagin. Este principio permitió asentar la teoría del control óptimo sobre un

riguroso fundamento matemático y abrió amplias posibilidades para su aplicación práctica en la rama de los sistemas de control automático. Sin embargo, en el presente libro no disponemos de la posibilidad de profundizar la teoría general del control automático basada en el principio del máximo y en este sentido rogamos refiéranse a la literatura al efecto.

Otro modo de tratar el cálculo de los procesos óptimos fue desarrollado por el matemático norteamericano R. Bellman y recibió la denominación de *programación dinámica.* El método de programación dinámica pone en manos del ingeniero un procedimiento de cálculo eficaz para resolver el problema de la optimización del control y está bien adaptado para el empleo de las calculadoras numéricas. Examinemos más detalladamente este método.

10-2. METODO DE PROGRAMACION DINAMICA

a) Forma discreta (discontinua) del problema variacional

Las dificultades examinadas de resolver el problema variacional se superan utilizando métodos de cálculo eficaces uno de los cuales es el método de programación dinámica. Este brinda la posibilidad de hallar el control óptimo en los problemas de pasos múltiples. No obstante, puede utilizarse también para resolver los problemas variacionales si éstos se presentan en forma discreta.

Recurriendo al teorema 10-1, enunciemos el problema variacional de la forma siguiente: hallar para el objeto descrito por la ecuación diferencial (10-1) el control $u(t)$ del dominio de controles admisibles U que minimiza la funcional

$$J(u) = \int_0^T Q[x(t), u(t)\,dt], \tag{10-16}$$

donde

$$Q(x, u) = Q_1(x, u) + \lambda H(x, u). \tag{10-17}$$

Obtendremos la forma discreta de notación de este problema si realizamos la elección del control $u(t)$ solamente en los momentos discretos $t = k\delta$, $k = 0, 1, \ldots, n-1$, donde $\delta = T/n$. Con esto, en lugar de la función $x(t)$ y $u(t)$ vamos a examinar las secuencias

$$x_k = x(t)\,|_{t=k\delta}, \qquad u_k = u(t)\,|_{t=k\delta}.$$

Substituyendo la derivada $\dot{x} = dx/dt$ en la ecuación (10-1) por la relación de los incrementos $(x_{k+1} - x_k)/\delta$ en lugar de la ecuación diferencial (10-1) obtenemos la ecuación de diferencias finitas

$$\frac{x_{k+1} - x_k}{\delta} = g(x_k, u_k). \tag{10-18}$$

De ordinario, esta ecuación se escribe en forma más cómoda resolviéndola con respecto a x_{k+1}:

$$x_{k+1} = x_k + g(x_k, u_k)\delta; \quad k = 0, 1, \ldots, n-1, \quad x_0 = c. \quad (10\text{-}19)$$

En este caso la integral (10-16) se sustituye por la suma

$$J_n(u) = \sum_{k=0}^{n-1} Q(x_k, u_k)\delta, \quad (10\text{-}20)$$

donde por u se entiende la secuencia de ecuaciones empleadas

$$u = (u_0, \ldots, u_{n-1}). \quad (10\text{-}21)$$

Ahora el problema consiste en elegir los controles $u_0, u_1, \ldots$ $\ldots, u_{n-1}$ que garantizan el valor mínimo de la suma (10-21).

En muchos problemas de control resulta conveniente considerar $\delta = 1$. Es particularmente cómodo hacerlo en los casos en que el proceso se divide en forma natural en pasos separados y, con esto, en los límites de cada paso el control $u(t)$ se mantiene invariable. Con esto llegamos al proceso de control de pasos múltiples estudiado en el capítulo 6 en el que x_k y u_k significan el estado del objeto y el control utilizado al comienzo de cada paso. Designando en este caso

$$x_k + g(x_k, u_k) = T(x_k, u_k), \quad (10\text{-}22)$$

escribiremos la ecuación (10-19) en la forma

$$x_{k+1} = T(x_k, u_k); \quad k = 0, 1, \ldots, n-1; \quad x_0 = c, \quad (10\text{-}23)$$

que coincide con la expresión (6-52) la cual define la transformación del estado del objeto durante un paso en el proceso de control de pasos múltiples. La suma (10-20) con $\delta = 1$ toma la forma de

$$J_n(u) = \sum_{k=0}^{n-1} Q(x_k, u_k), \quad (10\text{-}24)$$

que coincide con la expresión (6-54) para el criterio de calidad en el proceso de control de pasos múltiples.

En lo adelante vamos a estudiar el método de programación dinámica con arreglo a las relaciones (10-23) y (10-24), o sea, conforme al proceso de control de pasos múltiples. No obstante, la relación que hemos determinado entre las formulaciones matemáticas del problema variacional de forma discreta y el proceso de control de pasos múltiples, permite asimismo extender los resultados recibidos a la resolución de los problemas variacionales.

b) Relaciones recurrentes del método de programación dinámica

La optimización del control de un proceso de n-ésimo paso consiste en hallar la secuencia de controles $u_0, u_1, \ldots, u_{n-1}$, con la cual el criterio de calidad $J_n(u)$ toma el valor mínimo. Este

valor mínimo de dicho criterio de calidad de control del proceso de n-ésimo paso dependerá solamente del estado inicial x_0 y puede designársele como $f_n(x_0)$. Por definición tenemos:

$$f_n(x_0) = \min_{u_0} \min_{u_1} \dots \min_{u_{n-1}} [Q(x_0, u_0) + Q(x_1, u_1) + \dots$$
$$\dots + Q(x_{n-1}, u_{n-1})]. \quad (10\text{-}25)$$

Notemos que el primer sumando de esta expresión, $Q(x_0, u_0)$, depende solamente del control u_0 mientras que los demás sumandos dependen tanto de u_0 como de los controles en otros pasos. Así pues, $Q(x_1, u_1)$ depende de u_1, pero depende asimismo de u_0 ya que $x_1 = T(x_0, u_0)$. La situación de los sumandos restantes es análoga. Por eso, la expresión (10-25) puede escribirse en la forma

$$f_n(x_0) = \min_{u_0} \{Q(x_0, u_0) + \min_{u_1} \dots \min_{u_{n-1}} \times$$
$$\times [Q(x_1, u_1) + \dots + Q(x_{n-1}, u_{n-1})]\}. \quad (10\text{-}26)$$

Notemos más adelante que la expresión

$$\min_{u_1} \dots \min_{u_{n-1}} [Q(x_1, u_1) + \dots + Q(x_{n-1}, u_{n-1})] \quad (10\text{-}27)$$

constituye el valor mínimo del criterio de calidad de control $(n-1)$ del proceso de paso que tiene el estado inicial x_1. De acuerdo con la definición (10-25) podemos designar esta magnitud por $f_{n-1}(x_1)$. De este modo, obtenemos

$$f_n(x_0) = \min_{u_0} [Q(x_0, u_0) + f_{n-1}(x_1)]. \quad (10\text{-}28)$$

Estos razonamientos pueden repetirse si se examina el proceso de paso $(n-1)$ que comienza por el estado inicial x_1. El valor mínimo del criterio de calidad de control para este caso es

$$[f_{n-1}(x_1) = \min_{u_1} [Q(x_1, u_1) + f_{n-2}(x_2)]; \quad (10\text{-}29)$$

Al continuar estos razonamientos obtenemos una expresión análoga para el proceso de paso $n - l$ que comienza por el estado x_l:

$$f_{n-l}(x_l) = \min_{u_l} [Q(x_l, u_l) + f_{n-(l+1)}(x_{l+1})]. \quad (10\text{-}30)$$

La ecuación (10-30), que frecuentemente se denomina ecuación de Bellman, constituye una relación recurrente que permite determinar consecutivamente el control óptimo en cada paso del proceso controlado.

La propia idea de optimizar el control por separado en cada paso, si es difícil optimizar de una vez todo el proceso entero, no es original y se utiliza ampliamente en la práctica. No obstante, con frecuencia no se tiene en cuenta que la optimización de cada paso no significa aún la optimización de todo el proceso. Así pues, el sacrificar una figura en una partida de ajedrez nunca es venta-

joso desde el punto de vista de una jugada aislada, pero puede serlo considerando toda la partida. El gasto de recursos para la amortización no puede ser beneficioso desde el punto de vista de la coyuntura en un momento determinado, pero es beneficioso examinando el trabajo de la empresa durante un período prolongado.

La peculiaridad del método de programación dinámica consiste en que reune la sencillez de resolución del problema de optimizar el control en un paso separado con la previsión de que toma en cuenta las consecuencias más lejanas de este paso.

En el método de programación dinámica la elección del control en un paso separado se realiza no desde el punto de vista de los intereses del paso dado que se manifiestan en la minimización de las pérdidas en dicho paso, es decir, la magnitud $Q(x_l, u_l)$, sino desde el punto de vista de los intereses de todo el proceso que se traducen a la minimización de las pérdidas sumarias $Q(x_l, u_l) + f_{n-(l+1)}(x_{l+1})$ en todos los pasos siguientes. De aquí se deriva la propiedad fundamental del proceso óptimo consistente en que *sean cuales sean el estado y el control iniciales, los controles siguientes deben ser óptimos respecto al estado que es consecuencia de la utilización del primer control.*

De la propiedad fundamental del control óptimo se desprende que la optimización del mismo para una etapa arbitraria de un proceso de pasos múltiples consiste solamente en la elección de los controles siguientes. Por esto suele ser más cómodo tomar en consideración no aquellos pasos que ya se han dado sino los que quedan por dar para llevar el proceso a su estado final. Desde este punto de vista es conveniente escribir la ecuación (10-30) en otra forma.

En la expresión (10-30) la magnitud $n - l$ significa el número de pasos hasta el final del proceso. Designemos esta magnitud por k. Entonces, vamos a designar las magnitudes $x_l = x_{n-k}$ y $u_l = u_{n-k}$ sencillamente por x y u. Estas van a significar el estado del objeto y el control utilizado k pasos antes del final del proceso. Designemos por x' el estado siguiente, es decir, al cual el objeto pasa del estado x al emplear el control u. Esto va a ser $x_{n-(l+1)}$ en las designaciones anteriores. Con esto la ecuación (10-23) se escribirá en la forma

$$x' = T(x, u), \qquad (10\text{-}31)$$

y la relación recurrente (10-30) tomará la forma:

$$f_k(x) = \min_u [Q(x, u) + f_{k-1}(x')]. \qquad (10\text{-}32)$$

c) Aspectos de cálculo de la programación dinámica

La programación dinámica es el método numérico de resolución del problema de optimización del control y por eso va acompañada de cálculos bastante complejos. Pero nosotros no vamos

a dar demasiada importancia a esta peculiaridad suponiendo que los cómputos necesarios se realizan con calculadoras electrónicas.

La determinación del control óptimo en el k-ésimo paso arbitrario, contado a partir del momento final del proceso, se encuentra fundándose en las relaciones (10-31) y (10-32), las que para más comodidad conviene escribir en forma un tanto diferente al efectuar los cálculos numéricos. Designemos por $F_k(x, u)$ la magnitud del criterio de calidad de control de un proceso de k-ésimo paso con el control óptimo en los últimos $k-1$ pasos y el control arbitrario en el paso inicial. Entonces, la relación (10-32) puede escribirse en la forma

$$F_k(x, u) = Q(x, u) + f_{k-1}(x'); \qquad (10\text{-}33)$$

$$f_k(x) = \min_u F_k(x, u), \qquad (10\text{-}34)$$

donde x' se determina por (10-31).

Las variables x y u pueden tomar conjuntos finitos de valores o cambiar continuamente en algunas gamas. En el último caso efectuaremos la discretización de estas variables separando de ellas los conjuntos finitos de valores en lo posible equidistantes. Luego, en ambos casos las variables x y u pueden considerarse como elementos de los conjuntos finitos $X = \{x^{(1)}, \ldots, x^{(m)}\}$ y $U = \{u^{(1)}, \ldots, u^{r}\}$. Es cómodo calcular por anticipado el valor de $Q(x, u)$ y presentarlo en forma de tabla con el producto directo de los conjuntos $X \times U$. Esta tabla deberá conservarse en la memoria de la calculadora electrónica.

Los cálculos por las fórmulas (10-33) y (10-34) se realizan llenando la tabla 10-1. En las dos primeras columnas de la tabla están enumeradas diversas combinaciones de $x \in X$ y $u \in U$. Las demás columnas se llenan calculando las magnitudes correspondientes. En las últimas columnas de la tabla, para cada $x \in X$ se

Tabla 10-1

Tabla de cálculo por el método de programación dinámica

x	u	$x' = T(x, u)$	$Q(x, u)$
$x^{(1)}$	$u^{(1)}$	$T(x^{(1)}, u^{(1)})$	$Q(x^{(1)}, u^{(1)})$
	. . .	. . .	. . .
	$u^{(r)}$	$T(x^{(1)}, u^{(r)})$	$Q(x^{(1)}, u^{(r)})$
$x^{(2)}$	$u^{(1)}$	$T(x^{(2)}, u^{(1)})$	$Q(x^{(2)}, u^{(1)})$
	. . .	. . .	. . .
	$u^{(r)}$	$T(x^{(2)}, u^{(r)})$	$Q(x^{(2)}, u^{(r)})$
$x^{(3)}$	. . .	. . .	. . .

x	$f_{k-1}(x')$	$F_k(x, u)$	$f_k(x)$	u^*
$x^{(1)}$	$f_{k-1}[T(x^{(1)}, u^{(1)})]$ . . . $f_{k-1}[T(x^{(1)}, u^{(r)})]$	$F_k(x^{(1)}, u^{(1)})$ . . . $F_k(x^{(1)}, u^{(r)})$	$F_k(x^{(1)})$	u_1^*
$x^{(2)}$	$f_{k-1}[T(x^{(2)}, u^{(1)})]$ . . . $f_{k-1}[T(x^{(2)}, u^{(r)})]$	$F_k(x^{(2)}, u^{(1)})$ . . . $F_k(x^{(2)}, u^{(r)})$	$F_k(x^{(2)})$	u_2^*
$x^{(3)}$	. . .	. . .	. . .	. . .

escribe el valor mínimo de $F_k(x, u)$, es decir, se determinan $f_k(x)$ y el control óptimo u^*.

Es necesario realizar cálculos similares para cada paso del proceso de pasos múltiples. En este caso hay que tomar en consideración que para determinar $f_k(x)$ hace falta disponer previamente de la tabla para $f_{k-1}(x)$ ya que con su auxilio se hallan los valores de $f_{k-1}(x')$ al llenar la tabla 10-1. Por tanto, los valores de $f_{k-1}(x)$ deben calcularse antes que los valores de $f_k(x)$. De aquí se desprende que es preciso comenzar a calcular las funciones $f_k(x)$ desde el último paso del proceso de pasos múltiples. Notando que $f_0(x) = 0$, como etapa inicial del cálculo debe tomarse $k = 1$, para la cual

$$F_1(x, u) = Q(x, u); \qquad f_1(x) = \min_u Q(x, u). \tag{10-35}$$

En lo sucesivo el cálculo se realiza del modo habitual para $k = 2, 3, \ldots, n$.

Después de efectuar los cálculos y componer las tablas para $f_k(x)$ y u^* con $k = 1, \ldots, n$, puede comenzarse la búsqueda del control óptimo para todo el proceso con las condiciones iniciales dadas x_0. Por la tabla para $f_n(x)$ se halla u_0^* correspondiente a la x_0 dada y se calcula $x_1 = T(x_0, u_0^*)$. Más adelante, por la tabla para $f_{n-1}(x)$ se halla u_1^* correspondiente a la x_1 hallada y se calcula $x_2 = T(x_1, u_1^*)$, etc. Como resultado obtenemos el control óptimo $u^* = (u_0^*, u_1^*, \ldots, u_{n-1}^*)$.

En caso de que u sea la magnitud continua, $F_k(x^{(i)}, u)$ será función continua de u. Pero la tabla 10-1 da únicamente los valores discretos de $F_k(x^{(i)}, u^{(j)})$, $j = 1, \ldots, r$ de esta función. Con esto puede obtenerse el valor mínimo de $F_k(x^{(i)}, u)$ en los valores intermedios de u. En este caso deben determinarse las magnitudes u^*

y $f_h(x^{(i)})$ utilizando las fórmulas de interpolación. También es necesario emplear las fórmulas de interpolación para determinar el control óptimo u_k^* en los casos en que los valores de $x_k = T(x_{k-1}, u_{k-1})$ van a encontrarse entre los valores de x que forman parte de la tabla 10-1.

Ejemplo 10-1. Problema de la toma de altura. Un avión vuela con velocidad de v_0 a una altura de h_0. Es necesario cambiar su velocidad hasta v_1 y su altura hasta h_1 de modo que el gasto de combustible para este cambio sea mínimo.

Vamos a representar el proceso de cambio de v y h en el plano (v, h). Realicemos la discretización de las variables dividiendo cada gama de cambio en cuatro intervalos. Entonces, los estados discretos del objeto controlado van a representarse como nudos de la retícula representada en la figura 10-2. Consideremos que en cada nudo de la retícula sólo pueden emplearse dos controles:

$u_i = 0$, solamente el cambio de la velocidad v;

$u_i = 1$, solamente el cambio de la altura h.

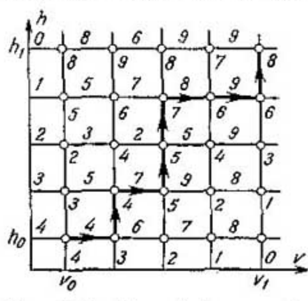


Fig. 10-2. Trayectoria en el problema de la toma de altura

De este modo, el conjunto de controles admisibles será el conjunto $U = \{0, 1\}$.

En la figura 10-2 está representada una de las trayectorias posibles correspondiente al control $u = (01011001)$. Para estimar esta trayectoria hay que conocer el gasto de combustible en cada paso. Esto será precisamente la función de objetivo $Q(x, u)$. Presentaremos los valores de $Q(x, u)$ en forma de números convencionales con los cuales en la figura 10-2 está marcada cada una de las transiciones posibles. Para la trayectoria representada en la figura 10-2 el gasto sumario de combustible que representa los valores del criterio de eficacia de control es igual a: $4 + 4 + 7 + 5 + 7 + 8 + 9 + 8 = 52$.

Para representar el proceso estudiado como de pasos múltiples, introduzcamos el método conveniente para describir los estados del objeto controlado. Señalemos los valores discretos v con los números del 0 al 4 comenzando por el valor final. Obremos del mismo modo respecto a h. Entonces, x_{ij} va a significar el estado para $v = i$ y $h = j$ a partir del cual quedan por dar $i + j$ pasos hasta el final del proceso.

Tabla 10-2

Cálculo del control óptimo en el último paso

x	u	x'	$Q(x, u)$	$f_1(x)$	u^*
x_{10}	0 1	x_{00} —	9 —	9	0
x_{01}	0 1	— x_{00}	— 8	8	1

Designemos por X_k el conjunto de estados con los que el proceso termina durante k pasos. Formarán parte de este conjunto todas aquellas x_{ij} para las cuales $i + j = k$. Considerando que $k = 0, 1, 2, \ldots,$ obtendremos

$$X_0 = \{x_{00}\}; \qquad X_1 = \{x_{10}, x_{01}\}; \qquad X_2 = \{x_{20}, x_{11}, x_{02}\}, \qquad \text{etc.}$$

Ahora puede iniciarse la resolución del problema. Para $k = 1$ tenemos:

$$X_1 = \{x_{10}, x_{01}\}; \quad F_1(x, u) = Q(x, u); \quad f_1(x) = \min_u F_1(x, u).$$

Utilicemos estas relaciones para componer la tabla 10-2 de estructura similar a la tabla 10-1.

Tabla 10-3

Cálculo del control óptimo en el penúltimo paso

x	u	x'	$Q(x, u)$	$f_1(x')$	$F_2(x, u)$	$f_2(x)$	u^*
x_{20}	0	x_{10}	9	9	18	18	0
	1	—	—	—	—		
x_{11}	0	x_{01}	9	8	17	16	1
	1	x_{10}	7	9	16		
x_{03}	0	—	—	—	—	14	1
	1	x_{01}	6	8	14		

Tabla 10-4

Cálculo del control óptimo a tres pasos hasta el final

x	u	x'	$Q(x, u)$	$f_2(x')$	$F_3(x, u)$	$f_3(x)$	u^*
x_{30}	0	x_{20}	6	18	24	24	0
	1	—	—	—	—		
x_{21}	0	x_{11}	8	16	24	24	0
	1	x_{20}	8	18	26		
x_{12}	0	x_{02}	9	14	23	22	1
	1	x_{11}	6	16	22		
x_{03}	0	—	—	—	—	17	1
	1	x_{02}	3	14	17		

Con $k = 2$, tenemos:

$$X_2 = \{x_{20},\ x_{11},\ x_{02}\};$$
$$F_2(x, u) = Q(x, u) + f_1(x'), \quad f_2(x) = \min_u F_2(x, u).$$

Utilicemos estas correlaciones para componer la tabla 10-3. Los cálculos análogos se realizan para $k = 3, 4 \ldots$

En el ejemplo examinado los datos contenidos en las tablas 10-2—10-4 pueden representarse directamente en el plano (v, h), señalando los valores de $f_k(x)$ en forma del número en el nudo correspondiente de la retícula, y de u^* con una flecha dirigida al nudo siguiente como se muestra en la figura 10-3. Después que se han determinado todos los valores de $f_k(x)$ y u^* para todos los nudos de la retícula, hallamos la trayectoria óptima moviéndonos desde el nudo inicial en la dirección de las flechas. El control óptimo será $u_* = (11000110)$ al cual corresponde $I_n(u_*) = 37$.

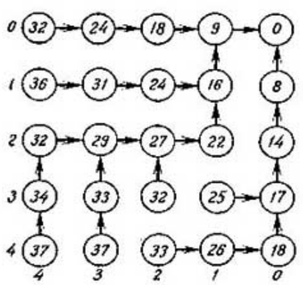


Fig. 10-3. Trayectoria óptima en el problema de la toma de altura

El ejemplo examinado ilustra perfectamente las ventajas del método de programación dinámica. Dicho método, en primera, elimina el problema del mínimo absoluto y relativo, ya que por el mismo proceso de cálculos está claro que siempre se halla el mínimo absoluto. En segunda, las limitaciones del tipo $|u_i| \leqslant M_i$ que son un serio obstáculo para el empleo de los métodos variacionales sólo facilitan el proceso de cálculo por el método de programación dinámica ya que reducen el dominio de controles admisibles U. Por último, el método de programación dinámica simplifica incomparablemente la búsqueda de la solución óptima con respecto al método de simple selección de variantes. Esto puede ilustrarse mediante un cálculo elemental.

Si en cada paso pueden haber r controles diferentes, en caso de selección directa cada uno de ellos debe examinarse en combinación con todos los controles posibles en todos los pasos restantes, lo que en el proceso del n-ésimo paso proporciona r^n variantes. Incluso para un problema muy sencillo de pasos múltiples con $r = 10$ y $n = 10$ se obtiene un número enorme 10^{10} de variantes.

En el método de programación dinámica, al elegir el control en algún paso para el estado x, no se examinan todas las continuaciones posibles sino solamente aquellas que corresponden a las continuaciones óptimas del estado $x' = T(x, u)$. Esto permite excluir del examen una enorme cantidad de variantes que no son de interés. De este modo, si en cada paso pueden haber m estados y para cada estado, r controles, el número total de variantes que se requiere examinar será rmn o r^2n si consideramos que $r = m$, es decir, que existe un control que convierte cada estado en cualquier otro nuevo. Con $r = 10$ y $n = 10$ se obtienen en total 10^3 variantes.

No obstante, pese a sus grandes ventajas, el método de programación dinámica tiene sus inconvenientes. Estos consisten en que para hallar el control óptimo con cierto estado inicial del objeto es necesario ir desde el final hasta el principio del proceso, determinando en cada paso los controles óptimos para todos los estados posibles del objeto en este paso. Y hasta el último momento permanece desconocido cuál será el control óptimo para el estado inicial dado. Con esto, en definitiva resulta que la mayor parte de la labor de cálculo se realizó inútilmente ya que los resultados de la determinación del control óptimo no se utilizan para los estados que no se encuentran en la trayectoria óptima. En este sentido tiene ventajas el principio de máximo mencionado anteriormente que hace posible construir una trayectoria óptima tomada por separado.

d) Control del estado final

Como se indicó en el capítulo 6 en algunos problemas no interesa el carácter del movimiento del objeto durante el proceso de control y sólo importa el estado x_n al que pasa el objeto cuando termina este proceso. En este caso servirá de criterio de calidad de control el valor de la función de objetivo al final del proceso de control, o sea, la magnitud

$$q = q(x_n). \tag{10-36}$$

Designemos como antes por x y u el estado del objeto y el mando empleado por k pasos antes del final del proceso, y por $x' = T(x, u)$ el estado siguiente, es decir, el estado en $k - 1$ pasos antes del final del proceso. El estado x_n y, vale decir, la función de objetivo $q(x_n)$ se determinan tanto por el valor inicial de x como por todos los controles ulteriores.

Para obtener la relación recurrente que determina el control óptimo, designemos por $F_k(x, u)$ el valor de la función de objetivo $q(x_n)$ en k pasos hasta el final del proceso con el estado inicial x, el control arbitrario u en el paso inicial y el control óptimo en los $k - 1$ pasos restantes, y por $f_k(x)$ el valor $q(x_n)$ con el control óptimo en todos los k pasos hasta el final del proceso, de suerte que

$$f_k(x) = \min_u F_k(x, u). \tag{10-37}$$

El valor $F_k(x, u)$ puede obtenerse gracias a las consideraciones siguientes. Bajo la acción del control arbitrario u el objeto controlado en el paso inicial se pasa del estado x al estado $x' = T(x, u)$. Para obtener $q(x_n) = F_k(x, u)$ es necesario emplear en los $k - 1$ pasos restantes el control óptimo lo que da $q(x_n) = f_{k-1}(x') = f_{k-1}[T(x, u)]$. Este será el valor de $F_k(x, u)$. De este modo,

$$F_k(x, u) = f_{k-1}[T(x, u)]. \tag{10-38}$$

Sustituyendo los valores $F_k(x, u)$ en (10-37), obtenemos

$$f_k(x) = \min_u f_{k-1}[T(x, u)]. \qquad (10\text{-}39)$$

La fórmula (10-39) es completamente análoga por su estructura a la fórmula (10-32). Aplicándole el método de cálculos ya descrito y tomando en consideración que

$$f_0(x) = q(x), \qquad (10\text{-}40)$$

puede hallarse el control óptimo para cualquier estado inicial del proceso de pasos múltiples.

e) Relación recurrente para los procesos de Márkov

Examinemos el método de programación dinámica para optimizar los procesos estocásticos en caso de que dicho proceso sea una cadena de Márkov controlada.

Como se mostró en el § 5-6 la cadena de Márkov puede ser presentada en forma de matriz de probabilidades de las transiciones $\mathbf{P} = \| p_{ij} \|$ de orden $L \times L$. Para que la cadena de Márkov sea controlada debe existir la posibilidad de cambiar las probabilidades de las transiciones p_{ij} mediante ingerencia del exterior. Supongamos que hay la posibilidad de llevar a cabo el proceso de Márkov por N métodos diferentes, con esto, al método k corresponde la matriz de probabilidades de las transiciones $P_k = \| p_{ij}^{(k)} \|$, $k = 1, \ldots, N$. Llamaremos estrategia a cada método de realización del proceso de Márkov.

En lo sucesivo vamos a considerar que tenemos la posibilidad de estimar cada método de realización del proceso, presentando la matriz de pérdidas o, lo que es más cómodo en el caso dado, la matriz de ganancias $\mathbf{R}_k = \| r_{ij}^{(k)} \|$. Las magnitudes $r_{ij}^{(k)}$, $i, j = 1, \ldots, L$; $k = 1, \ldots, N$ expresan la ganancia en un paso al pasar el proceso del estado i al estado j en caso de utilizar la estrategia k.

Ejemplo 10-2. La fábrica de televisores del ejemplo 5-9, hallándose en el estado 1, puede aumentar la demanda organizando su propaganda lo que, no obstante, exige gastos adicionales y disminuye los ingresos. Si la fábrica se encuentra en el estado 2, ella puede acelerar la transición al estado 1 aumentando los gastos de investigación. Entonces, la estrategia 1 consiste en no sufrir gastos de propaganda e investigación y la estrategia 2, en efectuar dichos gastos. Las matrices de probabilidades de transiciones y de ingresos para estas estrategias pueden ser las siguientes:

$$\mathbf{P}_1 = \begin{Vmatrix} 0{,}5 & 0{,}5 \\ 0{,}4 & 0{,}6 \end{Vmatrix}; \qquad \mathbf{R}_1 = \begin{Vmatrix} 9 & 3 \\ 3 & -7 \end{Vmatrix};$$

$$\mathbf{P}_2 = \begin{Vmatrix} 0{,}8 & 0{,}2 \\ 0{,}7 & 0{,}3 \end{Vmatrix}; \qquad \mathbf{R}_2 = \begin{Vmatrix} 4 & 4 \\ 1 & -19 \end{Vmatrix}.$$

Planteemos el problema de hallar el método de selección de la estrategia que garantice la obtención del ingreso máximo en n pasos del proceso de Márkov controlado.

Designemos por $f_i(n)$ la magnitud máxima del ingreso en n pasos que comienzan por el estado i. Supongamos que del estado i se ha dado un paso al estado j con el ingreso $r_{ij}^{(k)}$ y las restantes $n-1$ transiciones se han efectuado de manera óptima. Con esto el ingreso completo en n pasos será $r_{ij}^{(k)} + f_j(n-1)$.

No obstante, en realidad la primera transición del tipo (i, j) se realiza con una probabilidad $p_{ij}^{(k)}$. Por cuanto la primera transición es aleatoria, hay que tener en cuenta la posibilidad de pasar a todos los estados posibles $j \in \{1, \ldots, N\}$. Por eso el ingreso esperado será igual a:

$$F_i(n, k) = \sum_{j=1}^{N} p_{ij}^{(k)} [r_{ij}^{(k)} + f_j(n-1)] = q_i^{(k)} + \sum_{j=1}^{N} p_{ij}^{(k)} f_j(n-1), \quad (10\text{-}41)$$

donde la magnitud

$$q_i^{(k)} = \sum_{j=1}^{N} p_{ij}^{(k)} r_{ij}^{(k)} \quad (10.42)$$

brinda el ingreso en un paso desde el estado i con la estrategia k.

Forma parte de la expresión (10-41) la estrategia k empleada en el paso inicial y con la elección correspondiente de la cual podemos garantizar el ingreso máximo esperado $f_i(n)$ lo que nos conduce a elegir la relación recurrente para elegir la estrategia k en el primer paso del proceso de n pasos:

$$f_i(n) = \max_k \left[q_i^{(k)} + \sum_{j=1} p_{ij}^{(k)} f_j(n-1) \right]. \quad (10\text{-}43)$$

Es cómodo efectuar el cálculo aplicando la fórmula (10-43) por el método tabular análoga al examinado en los párrafos anteriores. Mostremos las tablas correspondientes (tablas 10-5 y 10-6) para el problema del ejemplo 10-2 para $n = 1$ y $n = 2$.

Tabla 10-5

$n = 1$

i	k	$p_{ij}^{(k)}$		$r_{ij}^{(k)}$		$q_i^{(k)}$	$f_i(1)$	k^*
		$j=1$	$j=2$	$j=1$	$j=2$			
1	1	0,5	0,5	9	3	6	6	1
	2	0,8	0,2	4	4	4		
2	1	0,4	0,6	3	−7	−3	−3	1
	2	0,7	0,3	1	−19	−5		

Tabla 10-6

$n = 2$

i	k	$q_i^{(k)}$	$p_{ij}^{(k)}$		$p_{ij}^{(k)} f_j(1)$		$F_i(2, k)$	$f_i(2)$	k^*
			$j=1$	$j=2$	$j=1$	$j=2$			
1	1	6	0,5	0,5	3	−1,5	7,5	8,2	2
	2	4	0,8	0,2	4,8	−0,6	8,2		
2	1	−3	0,4	0,6	2,4	−1,8	−2,4	−1,7	2
	2	−5	0,7	0,3	4,2	−0,9	−1,7		

Las tablas para $n > 2$ se compilan según el tipo de la tabla para $n = 2$.

PROBLEMAS PARA EL CAPITULO 10

10.1. Continúe la resolución del ejemplo 10-1 y calcule toda la trayectoria óptima.

10.2. Considerando la retícula de la figura 10-3 como grafo, halle la trayectoria óptima aplicando los métodos de búsqueda del camino mínimo en el grafo del § 2-2. Compare esta solutión con la solución por el método de programación dinámica.

10.3. Halle el control óptimo para el siguiente proceso de control de tres pasos de la magnitud escalar x. Al comienzo del proceso, x puede ser cualquier número entero de −10 a +10. En los diferentes pasos se realizan transformaciones del tipo $x' = T(x, u) = x + u$, con que los valores admisibles de u se determinan por los conjuntos:

$\{-1, 0, +1\}$ en el primer paso;
$\{-4, 0, +4\}$ en el segundo paso;
$\{-9, 0, +9\}$ en el tercer paso.

El objetivo del proceso es llevar en tres pasos la magnitud x hasta el valor deseando que se ha tomado igual a cero. Las pérdidas se estiman por el cuadrado de la desviación de x respecto a 0 al final del proceso, es decir, por la magnitud $q(x) = (x - 0)^2 = x^2$.

A NUESTROS LECTORES:

"Mir" edita libros soviéticos traducidos al español, inglés, francés, árabe y otros idiomas extranjeros. Entre ellos figuran las mejores obras de las distintas ramas de la ciencia y la técnica: manuales para los centros de enseñanza superior y escuelas tecnológicas; literatura sobre ciencias naturales y médicas. También se incluyen monografías, libros de divulgación científica y ciencia ficción.

Dirijan sus opiniones a la Editorial "Mir", I Rizhski per., 2, 129820, Moscú, I — 110, GSP, URSS.